

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 4, 2002, 11:29:22 ; Search time 4179.27 Seconds  
(without alignments)  
6378.264 Million cell updates/sec

Title: US-09-052-089a-8  
Perfect score: 1975  
Sequence: 1 GGCACGAGGTGGTGGTGGAC.....CAAAAAAAAAAAAAAAAAAAAA 1975

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapept 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

EST.\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estcov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrc:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1896	96.0	2696	11 AK012948	AK012948 Mus muscu
2	1886.2	95.5	1938	11 AK012786	AK012786 Mus muscu
3	671.2	34.0	887	9 AL560912	AL560912 AL560912
4	669.2	33.9	959	9 AL560947	AL560947 AL560947
5	641.2	32.5	702	10 B1151643	B1151643 602916189
6	634.2	32.1	687	10 BF658170	BF658170 maag6606
7	632.6	32.0	744	10 BF162255	BF162255 601770948
8	603	30.5	1061	10 BM464844	BM464844 AGENCOURT
9	568.2	28.8	608	10 B1696519	B1696519 603347473
10	539	27.3	589	10 BE334637	BE334637 us81912.Y
11	532.2	26.9	554	9 AA684194	AA684194 vm68d10.s
12	514.4	26.0	702	10 BG764118	BG764118 602737137
13	513.6	26.0	752	10 BG682548	BG682548 602624362
14	513.4	26.0	519	10 BE304189	BE304189 601086421
15	510.6	25.9	784	10 BG120736	BG120736 602347389
16	504.2	25.5	908	10 BG420765	BG420765 602448558
17	503.2	25.5	825	10 BE546959	BE546959 601071929

18	497.4	25.2	645	9 AL559248	AL559248
19	490.6	24.8	888	10 BE869186	BE869186
20	490.4	24.8	668	10 B1758046	B1758046 603030577
21	484.2	24.5	983	10 BE889971	BE889971 601512230
22	476.2	24.1	632	10 BF219707	BF219707 601296627
23	459.8	23.3	912	10 BG117979	BG117979 602351223
24	437.2	22.1	528	9 BB821649	BB821649
25	433	22.9	441	9 AA490717	AA490717 UT-M-BH3-
26	432.2	21.9	438	9 BB830459	BB830459
27	430	21.8	519	9 BB795850	BB795850
28	427.2	21.6	519	12 A2406449	A2406449 LM0175M08
29	424.4	21.5	547	9 BE031265	BE031265 129770 MA
30	422	21.4	450	9 BB831869	BB831869
31	407.8	20.6	453	9 BB830067	BB830067
32	406.2	20.6	412	9 BB835693	BB835693
33	401.4	20.3	423	9 BB779900	BB779900
34	399.8	20.2	403	9 A1428513	A1428513 mc37h05.x
35	388.4	19.7	391	9 AW823060	AW823060 ut79c12.x
36	385.6	19.5	393	9 BB837528	BB837528
37	373.6	18.9	571	10 B1341090	B1341090 368355 MA
38	367.6	18.6	769	10 BF038722	BF038722 601460641
39	354.8	18.0	932	9 AL529568	AL529568
40	339	17.2	585	9 AA910227	AA910227 ut79c12.Y
41	326.8	16.5	330	9 AA015242	AA015242 mh30c10.r
42	326.2	16.5	605	9 BB634734	BB634734
43	316	16.0	433	9 BB864250	BB864250
44	310.8	15.7	775	9 AL582241	AL582241
45	299.2	15.1	400	10 BE651125	BE651125 UT-M-BH3-

## ALIGNMENTS

RESULT	1	ALIGNMENTS
AK012948	2696 bp mRNA	Linear HTC 19-JAN-2002
LOCUS	Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length	
DEFINITION	enriched library, clone:2810054N23:TRAF-Interacting protein, full	
ACCESSION	AK012948	
VERSION	AK012948.1 GI:12850018	
KEYWORDS	HTC: CAP trapper.	
SOURCE	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (sites)	
TITLE	Carninci, P. and Hayashizaki, Y.	
JOURNAL	High-efficiency full-length cDNA cloning	
MEDLINE	Meth. Enzymol. 303, 19-44 (1999)	
PUBMED	99279253	
AUTHORS	2 (sites)	
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
MEDLINE	Genome Res. 10 (10), 1617-1630 (2000)	
PUBMED	20499374	
REFERENCE	11042159	
AUTHORS	3 (sites)	
TITLE	Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, D., Nishi, K., Kitsu, T., Tashiro, T., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format	
JOURNAL	sequencing pipeline with 384 multicapillary sequencer	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	

MEDLINE PUBMED REFERENCE	JOURNAL AUTHORS	TITLE
20530913 11076861 4 (sites)	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001) 5 (bases 1 to 2696)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arkawa,T., Balderelli,R., Bono,H., Brownstein,M., Bull,C., Carinici,P., Fukuda,S., Fukushima,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Humo,D., Imocanli,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kuhihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numatazi,R., Ono,M., Okazaki,Y., Okido,T., Owa,C., Quackebush,J., Salto,H., Salto,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,I., Tanaka,Y., Tejima,Y., Toyota,T., Yamamura,T., Yamanaka,I., Yasunishiki,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.	Submitted	Direct Submission Submitted (10-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
	Please visit our web site ( <a href="http://genome.gsc.riken.go.jp/">http://genome.gsc.riken.go.jp/</a> ) for further details.	CDS
	CDS library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGGACGATCCAGAGCTCGTGTGGTTTGGTTTGGTA 3'. cDNA was prepared by using trehalose thermo-acetylated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGGAGCTTCGAGTTAATTAAATGAATCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOUR.	gene
	Location/Qualifiers	FEATURES
	1..2696	SOURCE
/organism="Mus musculus"		
/strain="C57BL/6J"		
/db_xref="MGI:MGI:1896093"		
/db_xref="taxon:10090"		
/clone="2810054N23"		
/clone_1lb="RIKEN full-length enriched mouse cDNA library"		
/dev_stage="10, 11 days embryo"		
112..1524		
/gene="Tralp"		
112..1524		
/gene="Tralp"		
/note="TRAF-interacting protein data source:MCD, source key:MGI:1096377, evidence:ISS putative"		
/codon_start=1		
/protein_id="BA82567.1"		
/db_xref="GI:12850019"		
/translation="MPRALCTTCSDFSDHSDVVAIHGCHTEHLQCLIQMFETAPSR TCPCRDIOEKRTITINKLFPDLAQEEYNVDIAEFLKNELSDSVKASLSOKRERKDSQA ITDIPRLTERGNATVESLONALKAEMLCSTLKQMKPLEORODERTKREBNHRKL CKMTMOIELOLSORSSEVEEMRMDCVSANEOIAVCYSIKREYENLKERRKT GKMLADRKDIIVRSRKLTIUNTELDAKLELRASQNDISADDEITSLKKKITLOG TLSPTATNETVSRVLSPESPVEPMNPRILHPPEGEIDLDITFNTVPPTTGSGSGLCPKPKTLERRARSPMONVNLKVKHAKPSPOLSIGRCVGEIODEIELAACPLFIIN AVLGKQPNNTTASRCSSTDVVRIGFGIGLGRFTIOPDDTIIIRPVYSKAKSKOK		

		VRIKTVSSASDPKLDIFLCO"
	polyA_signal	2678..2683
	/note="putative"	
	polyA_site	2696
	/note="putative"	
BASE COUNT	699 a	676 c 711 g 610 t
ORIGIN		
Query Match	96.0%;	Score 1896; DB 11; Length 2696;
Best Local Similarity	99.0%;	Pred. No. 0;
Matches 1939;	Conservative 0;	Mismatches 15; Indels 4; Gaps 3.
OY	4	ACGAGGTGGCGTGCACCGAAATTTGAAGAACCAGCCGAGCGGTGCCGCTTC--CACCAAAC 61
DB	2	AAGTCGTGGGTGAGACGGAATAATTGAAGAACCCGAGCGGTGCGCTTGCGCCACAAC 61
OY	62	TGTGTGTGTCTCTGGGACGTGTGTTCCCTGGGCTGCTTGATGTGAGCATCATGCTCATTC 121
DB	62	TGTGTGTGTCTGGGTGAGAGTGTGTTCCCTGGGCTGCTTGATGTGAGCATCATGCTCATTC 121
OY	122	TCCTCTGTGCACATATCTCTCCGACTTCTTGATGATCACCCTGACGCTGGCTGCATCC 181
DB	122	GCGCTTGTGTGCATATCTCTCCGACTTCTTGATGATCACCCTGACGCTGGCTGCATCC 181
OY	182	ACTGTGGCCACACTTTTCATCTGCATATGCCATATCCAGTAGTGTGTAAGACAGCACAATGC 241
DB	182	ACTGTGGCCACACTTTTCATCTGCATATGCCATATCCAGTAGTGTGTAAGACAGCACAATGC 241
OY	242	GGACCTGCCCCACAGTGTAGATTCACAGTGTGGCAAANAAGCATATTTAAACAACCTTTTCT 301
DB	242	GGACCTGCCCCACAGTGTAGATTCACAGTGTGGCAAANAAGCATATTTAAACAACCTTTTCT 301
OY	302	TTGACCTGGCCAGGAAGAGGAATGTCTTGATGTGCACATTCATTAAGAATGAACCTGG 361
DB	302	TTGACCTGGCCAGGAAGAGGAATGTCTTGATGTGCACATTCATTAAGAATGAACCTGG 361
OY	362	ACAGCGTCAAAAGCTCAGCTTTTCCAGAAAAGACAGGAGAGAAACGGGACAGCCAGGCATTTA 421
DB	362	ACAGCGTCAAAAGCTCAGCTTTTCCAGAAAAGACAGGAGAGAAACGGGACAGCCAGGCATTTA 421
OY	422	TGCACACTCTACGGGACACCCCTGGAGAAGACGCAATGTACCGTGGAGTCCCTACAGAACG 481
DB	422	TGCACACTCTACGGGACACCCCTGGAGAAGACGCAATGTACCGTGGAGTCCCTACAGAACG 481
OY	482	CCTTAAACAGGACAGAGTGTGTCTTCCACCTGA AAAAACAAGATGAAGTCCCTGGAGC 541
DB	482	CCTTAAACAGGACAGAGTGTGTCTTCCACCTGA AAAAACAAGATGAAGTCCCTGGAGC 541
OY	542	ACGCGAGAGTATAGACCAACAAGCTCCGGAGAGAGGCCACCGACATCCAAGTCAAGTAGTA 601
DB	542	ACGCGAGAGTATAGACCAACAAGCTCCGGAGAGAGGCCACCGACATCCAAGTCAAGTAGTA 601
OY	602	AAACCATGAGGAGCAAAATTGAGCTCTACTTCAGAGCCAGCGTTCTGAGGTGAGAGATGA 661
DB	602	AAACCATGAGGAGCAAAATTGAGCTCTACTTCAGAGCCAGCGTTCTGAGGTGAGAGATGA 661
OY	662	TTTCGAGACATGGGTGTGGGACAGTCAAGCGCTGAGACAGTGGCTGTGTA CTGCTGCC 721
DB	662	TTTCGAGACATGGGTGTGGGACAGTCAAGCGCTGAGACAGTGGCTGTGTA CTGCTGCC 721
OY	722	TCGAAGAAAGATGATGAATTCGAAGGAAGCTCGGAAGGCCACAGGGGAACCTGGCTGACA 781
DB	722	TCGAAGAAAGATGATGAATTCGAAGGAAGCTCGGAAGGCCACAGGGGAACCTGGCTGACA 781
OY	782	GGTTGAAGAAGATTGTGTCTCTTAGAGGCAAGTTGAAGACTCTCAACACTGAGCTGG 841
DB	782	GGTTGAAGAAGATTGTGTCTCTTAGAGGCAAGTTGAAGACTCTCAACACTGAGCTGG 841
OY	842	ATTCAGGCCAAGTTAGAACTGAGGTACAGCCCAAGAGACTTACAAAGTCTGACACAGAGA 901
DB	842	ATTCAGGCCAAGTTAGAACTGAGGTACAGCCCAAGAGACTTACAAAGTCTGACACAGAGA 901

QY	902	TCACGACCCTTAAGAAAGAAAGTCTGATGATGATCTTCAGAGGAACCTTAAAGCCTGACCTCC	-GC6	960
Dp	902	TCACGACCCTTAAGAAAGAAAG	-CTGATGATCTTCAGAGGAACCTTAAAGCCTGACCTCCGCGC	960
QY	961	ACCAATGAGACGGTCAGACGCCCTCGTGTGTTTGGAGAGCCCGACCCCTGCTGAGAGTGTATGAC		1020
Dp	961	ACCAATGAGACGGTCAGACGCCCTCGTGTGTTTGGAGAGCCCGACCCCTGCTGAGAGTGTATGAC		1020
QY	1021	CCGAGGCTTCACACGACCACCTTCGGGTGATGAGATTTGATCTCAATTAACACCTTTGATGTA		1080
Dp	1021	CCGAGGCTTCACACGACCACCTTCGGGTGATGAGATTTGATCTCAATTAACACCTTTGATGTA		1080
QY	1081	AATACCCCTCCAAACCCAGACCTCTGGCTCCGACATGCTCCCAAGAAAGCTGTGCTGT		1140
Dp	1081	AATACCCCTCCAAACCCAGACCTCTGGCTCCGACATGCTCCCAAGAAAGCTGTGCTGT		1140
QY	1141	GAGAGGGCAGCGTCTCCCATGCAAGATGTCTCAAGAAAGTGCACAAAGTCTCCAAAGCGC		1200
Dp	1141	GAGAGGGCAGCGTCTCCCATGCAAGATGTCTCTCAAGAAAGTGCACAAAGTCTCCAAAGCGC		1200
QY	1201	GAGTCCACGCTCTCAGTGGGTGGCCAGCGATGTGTAGAGAGCTAGATGAGAGACTGGCT		1260
Dp	1201	GAGTCCACGAGCTCTCAGTGGGTGGCCAGCGATGTGTAGAGAGCTAGATGAGAGAACTGGCT		1260
QY	1261	GGTGCCTTCCTCTCTTCATTCGCGAATGCGTCTGGGTCAGAAACAGGCCCAACAGAGACC		1320
Dp	1261	GGTGCCTTCCTCTCTCTTCATTCGCGAATGCTGTCTGTGGTTCAGAAACAGGCCCAACAGAGACC		1320
QY	1321	ACAGCAGAAATCCCGAAGCAGCAGACAGATGTGTAAAGATAGGCTTTGATGGGCTTGGAGGA		1380
Dp	1321	ACAGCAGAAATCCCGTTCAGCAGACAGATGTGTAAAGATAGGCTTTGATGGGCTTGGAGGA		1380
QY	1381	CGAACAAATTCATCCAGCCTTAGGGACACAAACCATATCTCGACACAGTGCCTGTAAATGCC		1440
Dp	1381	CGAACAAATTCATCCAGCCTTAGGGACACAAACCATATCTCGACACAGTGCCTGTAAATGCC		1440
QY	1441	AAGGCCAAGAGTAAACAGAAAGTGAAGATTAAGACTGTAGTTCGCTCCGACGCCCAAG		1500
Dp	1441	AAGGCCAAGAGTAAACAGAAAGTGAAGATTAAGACTGTAGTTCGCTCCGACGCCCAAG		1500
QY	1501	CTGGATTAACCTTCTTATATGTCAGTGAAGCGGTGACCAAGTATGTTTGCAATTAATGTGGCCCA		1560
Dp	1501	CTGGATTAACCTTCTTATATGTCAGTGAAGCGGTGACCAAGTATGTTTGCAATTAATGTGGCCCA		1560
QY	1561	AGACCTGGCTAACCGGAAGTGTGTTTGGAAAGATGGCTCTCTTGGACCAAGTCCAAAGABAG		1620
Dp	1561	AGACCTGGCTAACCGGAAGTGTGTTTGGAAAGATGGCTCTCTTGGACCAAGTCCAAAGABAG		1620
QY	1621	ATGCCCAGAAAACACACTTCTCTGTTCACCTGCGCCCTACACACACTGGGAAGCCACAT		1680
Dp	1621	ATGCCCAGAAAACACACTTCTCTGTTCACCTGCGCCCTACACACACTGGGAAGCCACAT		1680
QY	1681	GACCAAGTTACTGTTCCGATCAGACAGAGGCCCTAATCCAGTGCAGAGGTTTTCCTATAGC		1740
Dp	1681	GACCAAGTTACTGTTCCGATCAGACAGAGGCCCTAATCCAGTGCAGAGGTTTTCCTATAGC		1740
QY	1741	TACAACCAAGGTGGCTGAGACTCCTTTGTTTATAGAAACAGGGTCAACATGACTCTAA		1800
Dp	1741	TACAACCAAGGTGGCTGAGACTCCTTTGTTTATAGAAACAGGGTCAACATGACTCTAA		1800
QY	1801	GTGGATGGAGTGTGTGAGAGATCTCTATGCAAGCTGGAGAGACCTTGCGCTTGAATCTCTGC		1860
Dp	1801	GTGGATGGAGTGTGTGAGAGATCTCTATGCAAGCTGGAGAGACCTTGCGCTTGAATCTCTGC		1860
QY	1861	CTGCGCTCAGCTTATTTCTTGAATTAATGCGGTGAGAGTGTATAGGAAAGTGTGGGGA		1920
Dp	1861	CTGCGCTCAGCTTATTTCTTGAATTAATGCGGTGAGAGTGTATAGGAAAGTGTGGGGA		1920
QY	1921	AGTTTTCTGTCTAAATTAAGAGGATCTTTCTTCAA	1958	
Dp	1921	AGTTTTCTGTCTAAATTAAGAGGATCTTTCTTCTTCA	1958	

RESULT	2
AK012786	
LOCUS	1938 bp mRNA linear HTC 19-JAN-2002
DEFINITION	Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2810021M06:trfA-interacting protein, full insert sequence.
ACCESSION	AK012786
VERSION	AK012786.1 GI:12849758
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (strain:C57BL/6J) 10, 11 days embryo cDNA to mRNA, clone:1ib:RIKEN full-length enriched mouse cDNA library clone:2810021M06.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (sites)
TITLE	Carninci, P. and Hayashizaki, Y.
JOURNAL	High-efficiency full-length cDNA cloning
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2 (sites)
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3 (sites)
AUTHORS	Shibata, K., Itoh, M., Atzawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplexed sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076661
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
MEDLINE	5 (bases 1 to 1938)
PUBMED	11076661
REFERENCE	Adachi, J., Atzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Aikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegata, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, Y., Kojima, Y., Konno, H., Kouda, M., Koyama, K., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 7.5 and subtraction to  
Rot = 37.5. Second strand cDNA was prepared with the primer adapter  
of sequence [5'  
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SolR.

FEATURES  
source

Location/Qualifiers

1. 1938

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="MGD:MG1:1898099"

/db\_xref="taxon:10090"

/clone="2810021M06"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="10, 11 days embryo"

95.1507

/gene="Talp"

95.1507

/gene="Talp"

/note="Talp-interacting protein"

data source:MGD, source key:MG1:1096377, evidence:ISS

putative"

/codon\_start=1

/protein\_id="BAB28469.1"

/db\_xref="GI:12849759"

/translation="MPIRALCTISDFPDHSRDVVAIHGHTFHLOCLIQMFETAPSR  
TCPCRIQVQKKTIINKLFPLDLOEBEENLVDAEIKNELDSVKOLSKDREKDSQA  
IIDLRLTLEKRAIVSLONALKAELCTLKKQKFLERODETRKQAEHRLKAT  
CKMTMEQIELLOSORSSEVEEMIRDMGVGSAVEQLAVCVSLKKEYENLEKARKAT  
GELADRLKIDLVSRSLKLTLELIDQAKLELRSAQDLASADELISLRKMLILOG  
TSLPAPNETVSRVLESPAPVEMNRLHPFGEIDINTFEDVNTPTPTQSGSQ  
HCLPKICLERASPMQNVLKKVHKVSKPESQLSLGGORCYGELDEFLAAGFPFLFIN  
AVIGOKOPNRTTASRCSSTDVVRIGFDLGRRTFIPDRITTIIRPVVSKASKSKOK  
VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 7.5 and subtraction to  
Rot = 37.5. Second strand cDNA was prepared with the primer adapter  
of sequence [5'  
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SolR.

FEATURES  
source

Location/Qualifiers

1. 1938

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="MGD:MG1:1898099"

/db\_xref="taxon:10090"

/clone="2810021M06"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="10, 11 days embryo"

95.1507

/gene="Talp"

95.1507

/gene="Talp"

/note="Talp-interacting protein"

data source:MGD, source key:MG1:1096377, evidence:ISS

putative"

/codon\_start=1

/protein\_id="BAB28469.1"

/db\_xref="GI:12849759"

/translation="MPIRALCTISDFPDHSRDVVAIHGHTFHLOCLIQMFETAPSR  
TCPCRIQVQKKTIINKLFPLDLOEBEENLVDAEIKNELDSVKOLSKDREKDSQA  
IIDLRLTLEKRAIVSLONALKAELCTLKKQKFLERODETRKQAEHRLKAT  
CKMTMEQIELLOSORSSEVEEMIRDMGVGSAVEQLAVCVSLKKEYENLEKARKAT  
GELADRLKIDLVSRSLKLTLELIDQAKLELRSAQDLASADELISLRKMLILOG  
TSLPAPNETVSRVLESPAPVEMNRLHPFGEIDINTFEDVNTPTPTQSGSQ  
HCLPKICLERASPMQNVLKKVHKVSKPESQLSLGGORCYGELDEFLAAGFPFLFIN  
AVIGOKOPNRTTASRCSSTDVVRIGFDLGRRTFIPDRITTIIRPVVSKASKSKOK  
VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 7.5 and subtraction to  
Rot = 37.5. Second strand cDNA was prepared with the primer adapter  
of sequence [5'  
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SolR.

FEATURES  
source

Location/Qualifiers

1. 1938

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="MGD:MG1:1898099"

/db\_xref="taxon:10090"

/clone="2810021M06"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="10, 11 days embryo"

95.1507

/gene="Talp"

95.1507

/gene="Talp"

/note="Talp-interacting protein"

data source:MGD, source key:MG1:1096377, evidence:ISS

putative"

/codon\_start=1

/protein\_id="BAB28469.1"

/db\_xref="GI:12849759"

/translation="MPIRALCTISDFPDHSRDVVAIHGHTFHLOCLIQMFETAPSR  
TCPCRIQVQKKTIINKLFPLDLOEBEENLVDAEIKNELDSVKOLSKDREKDSQA  
IIDLRLTLEKRAIVSLONALKAELCTLKKQKFLERODETRKQAEHRLKAT  
CKMTMEQIELLOSORSSEVEEMIRDMGVGSAVEQLAVCVSLKKEYENLEKARKAT  
GELADRLKIDLVSRSLKLTLELIDQAKLELRSAQDLASADELISLRKMLILOG  
TSLPAPNETVSRVLESPAPVEMNRLHPFGEIDINTFEDVNTPTPTQSGSQ  
HCLPKICLERASPMQNVLKKVHKVSKPESQLSLGGORCYGELDEFLAAGFPFLFIN  
AVIGOKOPNRTTASRCSSTDVVRIGFDLGRRTFIPDRITTIIRPVVSKASKSKOK  
VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 7.5 and subtraction to  
Rot = 37.5. Second strand cDNA was prepared with the primer adapter  
of sequence [5'  
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SolR.

FEATURES  
source

Location/Qualifiers

1. 1938

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="MGD:MG1:1898099"

/db\_xref="taxon:10090"

/clone="2810021M06"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="10, 11 days embryo"

95.1507

/gene="Talp"

95.1507

/gene="Talp"

/note="Talp-interacting protein"

data source:MGD, source key:MG1:1096377, evidence:ISS

putative"

/codon\_start=1

/protein\_id="BAB28469.1"

/db\_xref="GI:12849759"

/translation="MPIRALCTISDFPDHSRDVVAIHGHTFHLOCLIQMFETAPSR  
TCPCRIQVQKKTIINKLFPLDLOEBEENLVDAEIKNELDSVKOLSKDREKDSQA  
IIDLRLTLEKRAIVSLONALKAELCTLKKQKFLERODETRKQAEHRLKAT  
CKMTMEQIELLOSORSSEVEEMIRDMGVGSAVEQLAVCVSLKKEYENLEKARKAT  
GELADRLKIDLVSRSLKLTLELIDQAKLELRSAQDLASADELISLRKMLILOG  
TSLPAPNETVSRVLESPAPVEMNRLHPFGEIDINTFEDVNTPTPTQSGSQ  
HCLPKICLERASPMQNVLKKVHKVSKPESQLSLGGORCYGELDEFLAAGFPFLFIN  
AVIGOKOPNRTTASRCSSTDVVRIGFDLGRRTFIPDRITTIIRPVVSKASKSKOK  
VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5' GAGAGAGAGAGATCCAGAGAGCTCTTTTCTTTTCTTAA 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 7.5 and subtraction to  
Rot = 37.5. Second strand cDNA was prepared with the primer adapter  
of sequence [5'  
GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SolR.

FEATURES  
source

Location/Qualifiers

1. 1938

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="MGD:MG1:1898099"

/db\_xref="taxon:10090"

/clone="2810021M06"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="10, 11 days embryo"

95.1507

/gene="Talp"

95.1507

/gene="Talp"

/note="Talp-interacting protein"

data source:MGD, source key:MG1:1096377, evidence:ISS

putative"

/codon\_start=1

/protein\_id="BAB28469.1"

/db\_xref="GI:12849759"

/translation="MPIRALCTISDFPDHSRDVVAIHGHTFHLOCLIQMFETAPSR  
TCPCRIQVQKKTIINKLFPLDLOEBEENLVDAEIKNELDSVKOLSKDREKDSQA  
IIDLRLTLEKRAIVSLONALKAELCTLKKQKFLERODETRKQAEHRLKAT  
CKMTMEQIELLOSORSSEVEEMIRDMGVGSAVEQLAVCVSLKKEYENLEKARKAT  
GELADRLKIDLVSRSLKLTLELIDQAKLELRSAQDLASADELISLRKMLILOG  
TSLPAPNETVSRVLESPAPVEMNRLHPFGEIDINTFEDVNTPTPTQSGSQ  
HCLPKICLERASPMQNVLKKVHKVSKPESQLSLGGORCYGELDEFLAAGFPFLFIN  
AVIGOKOPNRTTASRCSSTDVVRIGFDLGRRTFIPDRITTIIRPVVSKASKSKOK  
VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSSASQPKLIDFLCO"

VRITVSS

|||||  
Db 1440 AGAAGTGAATAAAGACGTGAGTCTGCTCCAGACCCAAAGCTGATACCTTTAT 1499  
QY 1517 GTGAGTGAAGGTGACAGAGTGTGTTGCAATTAGTGGCCCAAGACCTGCTAACCG 1576  
Db 1500 GTGAGTGAAGGTGACAGAGTGTGTTGCAATTAGTGGCCCAAGACCTGCTAACCG 1559  
QY 1577 AAGTGTGTTGGAAGATGAGTCTCTTGAGACAGTCCAGAGAGATGCCAGAAACACA 1636  
Db 1560 AAGTGTGTTGGAAGATGAGTCTCTTGAGACAGTCCAGAGAGATGCCAGAAACACA 1619  
QY 1637 CTTCCTGTGTTCACTGCGCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTC 1696  
Db 1620 CTTCCTGTGTTCACTGCGCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTC 1679  
QY 1697 CGATCAGCAGGGGCTACTCTCAGTTGAGAGGTTTGTGTTAGCTTACACAGAGTGGC 1756  
Db 1680 CGATCAGCAGGGGCTACTCTCAGTTGAGAGGTTTGTGTTAGCTTACACAGAGTGGC 1739  
QY 1757 TGCAGCTCTTTGTTTATAGAACAGAGTTCACATTGACTTAAGTGAAGTGGAGTCTG 1816  
Db 1740 TGCAGCTCTTTGTTTATAGAACAGAGTTCACATTGACTTAAGTGAAGTGGAGTCTG 1799  
QY 1817 GAGATCTCTATGACAGCTGAGAGCCTGCGCTTGAACCTCTGCTCCACCTTATT 1876  
Db 1800 GAGATCTCTATGACAGCTGAGAGCCTGCGCTTGAACCTCTGCTCCACCTTATT 1859  
QY 1877 GCTTGAATTAAGGGAGAGTGTGATAGGAAAGTGGGAAAGTTTCTGTATAA 1936  
Db 1860 GCTTGAATTAAGGGAGAGTGTGATAGGAAAGTGGGAAAGTTTCTGTATAA 1919  
QY 1937 TAAAAAGGATCTTTCTT 1955  
Db 1920 TAAAAAGGATCTTTCTT 1938

RESULT 3  
AL560912 887 bp mRNA linear EST 16-FEB-2001  
LOCUS AL560912 LTI\_NFL010\_BC2 Homo sapiens cDNA clone CS0DL005Y108 5  
DEFINITION prime, mRNA sequence.  
ACCESSION AL560912  
VERSION AL560912.1 GI:12907828  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 101 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
source  
1. 887  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DL005Y108"  
/clone\_1id="LTI\_NFL010\_BC2"  
/sex="male"  
/tissue\_type="B cells from Burkitt lymphoma"  
/note="Vector: pCMVSPORT 6, Site\_1: NotI; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed by  
Life Technologies. Contact : Feng Liang Life Technologies,  
a division of Invitrogen 9600 Medical Center Drive  
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371  
Email : fliang@lifetech.com URL :

|||||  
BASE COUNT 251 a 219 c 245 g 171 t 1 others  
ORIGIN  
Query Match 34.0%; Score 671.2; DB 9; Length 887;  
Best local Similarity 87.2%; Pred. No. 1,4e-128;  
Matches 769; Conservative 1; Mismatches 109; Indels 3; Gaps 3;  
QY 96 CTTCAGTGCAGCCATCATGCTTATCTCTCTGTGACATATGCTCCGACTTCTTCA 155  
Db 4 CTTCAGTGCAGCCATCATGCTTATCTCTGTGACATATGCTCCGACTTCTTCA 63  
QY 156 TCACTCCGAGACCTGAGCTCCATCCACTGTGACACACTTTTCATGCAATGCTTAAT 215  
Db 64 TCACTCCGAGACCTGAGCTCCATCCACTGTGACACACTTTTCATGCAATGCTTAAT 123  
QY 216 CCACTGTTTGAAGACAGCACCACCAAGTGGACCTGCCACAGTGAATCCAGTTGCAA 275  
Db 124 TCACTGTTTGAAGACAGCACCACCAAGTGGACCTGCCACAGTGAATCCAGTTGCAA 183  
QY 276 AAAGACTTATTAACAACCTTTCTTGTGACCTGGCCAGGAAGAGATGCTTGA 335  
Db 184 AAAGACTTATTAACAACCTTTCTTGTGACCTGGCCAGGAAGAGATGCTTGA 243  
QY 336 TGCAGAAATTTCTTAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 395  
Db 244 TGCAGAAATTTCTTAAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 303  
QY 396 GGAGAAACGGGACAGCAGCCATTTATGACACTCTACGGGACACCTGGAAGAACAA 455  
Db 304 GGAGAAACGGGACAGCAGCCATTTATGACACTCTACGGGACACCTGGAAGAACAA 363  
QY 456 TGTACCGTGTGAGTCCCTACAGAACGCTTAAACAAGCAGAGTCTGTCTTCCACCT 515  
Db 364 TGTACCGTGTGAGTCCCTACAGAACGCTTAAACAAGCAGAGTCTGTCTTCCACCT 423  
QY 516 GAAAGAAACAGATGAAGTCTCTGAGCAGCGCAGAGTGAAGCAGCAAGCTCGGAGGA 575  
Db 424 GAAAGAAACAGATGAAGTCTCTGAGCAGCGCAGAGTGAAGCAGCAAGCTCGGAGGA 483  
QY 576 GGCCACCAGCTCAAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT 635  
Db 484 GGCCACCAGCTCAAGTGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT 543  
QY 636 CCAGCGTTCTGAGGTGAGAGATGATTCAGACATGGGTGGGACAGTACGGGTGGA 695  
Db 544 CCAGCGTTCTGAGGTGAGAGATGATTCAGACATGGGTGGGACAGTACGGGTGGA 603  
QY 696 GCAGCTGGCTGTGACTGCGCTCCCTCAAGAAAGATGAGATGAGAGAGTGC 755  
Db 604 ACAGCTGGCTGTGACTGCGCTCCCTCAAGAAAGATGAGATGAGAGAGTGC 663  
QY 756 GAAGGCCACAGGGAAGTGGCTGACAGGTTGA- GAAGATTTGGTCTCTAGAGCA 814  
Db 664 GAAGGCCACAGGGAAGTGGCTGACAGGTTGA- GAAGATTTGGTCTCTAGAGCA 723  
QY 815 AGTTGAAGACTCTCAACACTGAGCTGATGAGCCCACTTGAAGTGAAGTGAAGTGAAGT 874  
Db 724 AGTTGAAGACTCTCAACACTGAGCTGATGAGCCCACTTGAAGTGAAGTGAAGTGAAGT 783  
QY 875 AGGACTTACAAAGTGTGACAGCAGAGATACAGCAGCTTAAGAAAGTCTGATGATCTC 934  
Db 784 AGGACTTACAAAGTGTGACAGCAGAGATACAGCAGCTTAAGAAAGTCTGATGATCTC 842  
QY 935 CAGGAAACCTTGAAGCTGCTCC- GCGACCAATGAGACGCTC 975  
Db 843 CAGGAAACCTTGAAGCTGCTCCACAGTGGSCAGTGAAGTGC 884

RESULT 4  
AL560947 959 bp mRNA linear EST 16-FEB-2001  
LOCUS AL560947

[illegible]

D	b		477	GCTGTCCTCCACACTGTAAAAAGCAGATGAAGAATTACTTAAAGCACAGCAAGCATGACGCCA	536
Oy			561	AACAAGCTCGGGAGGAGCCCCACCAGACTCAAAGTGACAGATGAATAAACCATGAGCAAAATTGA	620
D	b		537	ACAAGACAAAGAGGAGGCCTGCCGGCTCAAGGACCAAGATGAAGACCATGAGCAGATTGA	596
Oy			621	GCTCCTTAATCCAGAGCCACGTTGTGTAGTAGTGAGAGATGATTCCAGACATGGGTGGG	680
D	b		597	GCTTCTACTCCAGAGCCAGCCGCTGTAGGTGGAGAGATGATCCGACATGGGTGGG	656
Oy			681	ACAGTCAGGGGGGAGCACCCTGGCTGTGTACTGTACGTCCGTCCTCAAGAAAGATGTAGAA	740
D	b		657	ACAGTCAGGGGGGAGCACCCTGGCTGTGTACTGTGTGTCTCAAGAAAAGATGTAGAA	716
Oy			741	TCTGAAGAAAGCTCGGAAGCCCACAGAGGAGMACTGGCTGCAGAGTTGGAMAAGCAATTTGGT	800
D	b		717	TCTAAAGAAGGACGSAAGCCCTCAGAGGGAGAGGTGGCTGACA - RCTGASGAAGATTTTGT	775
Oy			801	GTCTCTTAGGACCAAGTTGAGAATCTTCAACACTGAGCTGATTCAGAGCCAATTAGAACT	860
D	b		776	TTCCTCCAGAACCAAGTTTCAAGACAGTCACTGTGAATGTGATTCAGAGCCAATTTAGAACT	835
Oy			861	GAGGTCAAGCCCAAGAAAGACTTACAAAGTCTGTACCAAGAGATTCAGAGCCTTAAGAAAGA	920
D	b		836	GAGTCAAGCCCAAGAAAGACTTACAGAGTCTGTAMAAAGGAAATCATGACCTTGAAGAAAGA	895
Oy			921	GTCTATGATGCTCTCAGGGAACCTTGAGCGCTGCTTC - CGCGACCAATGAGAGCGTCAGCC	979
D	b		896	GCTAACGATGCTGCAAGAAACCTTGGAACTCTCCACAAAGTGGSCAGTGAAGACTGTGACCC	955
Oy			980	GCCT GCT 983	
D	b		956	GCCT GCT 959	
R	E	S	B11S1643		
L	I	CUS	B11S1643	702 bp mRNA linear EST 05-JUL-2001	
D	F	NITION	602916189P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5066423 5'		
A	C	SSION	mRNA sequence.		
V	E	RSION	B11S1643		
K	E	YWORDS	B11S1643.1 GI:14611644		
S	O	URCE	EST.		
O	R	GANISM	house mouse.		
			Mus musculus		
R	E	ERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
A	H	THORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
T	I	TLE	NIH-MGC http://mgc.nci.nih.gov/.		
J	O	URNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
C	O	MENT	Unpublished (1999) Contact: Robert Strauszberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov/">http://image.llnl.gov/</a> Plate: LLM11179 row: b column: 24 High quality sequence stop: 694.		
F	E	ATURES	Location/Qualifiers		
S	O	OURCE	1..702		
			/organism="Mus musculus"		
			/strain="CZECH II (fetal)"		
			/db_xref="taxon:10090"		
			/clone="IMAGE:5066423"		
			/clone_lib="NCI CGAP Lu29"		
			/tissue_type="spontaneous tumor, metastatic to mammary stem cell origin."		
			/lab_host="DH10B"		
			/note="Organ: Lung; Vector: pCMV-SPORT6; Site: 1; Salt:		





OY	496	GAGATGCTGTGGTCCACCCAAAAAACAATGAATTTCCTGGAGCAGCGGCAGATGAG	555
Db	481	GGATGCTGTGTGCCACCCTTAATAATACAGATGAAGTTCTCGAGCAGCGCATGATGAG	540
OY	556	ACCAACAAGACTCGGAGAGAGGAGCCACCGACTCAAGTCAGAATGAATAAACATGAGCAA	615
Db	541	ACCAACAAGACTCGTGAGAGAGGCCCCATCGACTCGATGCAAGATGAATAAACATGAGCAA	600
OY	616	ATTGAGCTCTTACTCTCAGAGCCAGCGCTTGAGAGTGAGAGATGATTCAGACATGGGT	675
Db	601	ATTGAGCTTCTACTCTCAAGAGCCACGCTGTGATGTGGAAGATGATGATGAGAAATGGGT	660
OY	676	GTGGGACAGTACGCGGTGGAGCAGCTG	702
Db	661	GTGGACAGTACGCGGTGGAGCATGTG	687
RESULT	7		
LOCUS	BF162255	744 bp	mRNA linear EST 30-OCT-2000
DEFINITION	60177094BP1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3990371 5'		
ACCESSION	BF162255		
VERSION	BF162255.1	GI:11042462	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs@email.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: LLM4201 row: o column: 12		
FEATURES	High quality sequence stop: 730.		
source	Location/Qualifiers		
	1..744		
	/organism="Mus musculus"		
	/strain="C2ECH II (ferral)"		
	/db_xref="taxon:10090"		
	/clone_image="3990371"		
	/clone_lib="NCI_CGAP_Lu29"		
	/tissue_type="spontaneous tumor, metastatic to mammary."		
	/stem_cell_origin="Stem cell origin."		
	/lab_host="DH10B"		
	/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"		
BASE COUNT	201 a 183 c 212 g 148 t		
ORIGIN			
Query Match	32.0%; Score 632.6; DB 10; Length 744;		
Best Local Similarity	96.3%; Pred. No. 1.4e-120;		
Matches 702; Conservative	0; Mismatches 19; Indels 8; Gaps 5;		
OY	24	ATTGGAAGAAACCGAGCGGTGGCG--GTTCACCAAACTGTCTGTCTGGCAGC	80
Db	1	ATTGGAAGAAACCGAGCGGTGGCGGTTCGCGCACCGAAACTGTCTGTCTGGCAGC	60
OY	81	TGGTTCCTGGCGCTGTTGAGTGAGGCATATCCATCTCTCTGTGGACATCTTG	140
Db	61	TGGTTCCTGGCGCTGTTGAGTGAGGCATATCCATCTCTCTGTGGACATCTTG	120





Db 240 CCAGCCCAAGCTGATACCTTCTTATGTACAGTGAACGGTGACAGAGTCATGTTGCAAT 299  
 QY 1551 TAGTGGCCCAAGACCTGCTTAACCGGAAGTGTGTAAGATGGCTCCTCTTGGACCG 1610  
 Db 300 TAGTGGCCCAAGACCTGCTTAACCGGAAGTGTGTAAGATGGCTCCTCTTGGACCG 358  
 QY 1611 TCAGAGAGATGCCCCAGAAAAACACACTTCCTGTTGACAGCCGCTGCACACACTGG 1670  
 Db 359 TCAGAGAGATGCCCCAGAAAAACACACTTCCTGTTGACAGCCGCTGCACACACTGG 418  
 QY 1671 G-AAGCCCAATGACCACTTACTCTTCCGATCAGAGAGGCGCTTCCAGTGGCAGGTT 1729  
 Db 419 GAAACCCACATGACCACTTACTCTTCCGATCAGAGAGGCGCTTCCAGTGGCAGGTT 478  
 QY 1720 TTGCTTAAGCTACCAACAGGTGTGCTGACCTCTTGTGTTTATAGAACAGAGGTAC 1789  
 Db 479 TTGCTTAAGCTACCAACAGGTGTGCTGACCTCTTGTGTTTATAGAACAGAGGTAC 538  
 QY 1750 ATTGACTTAAGTGGATGGAGTGTGAGATCCTATGCAAGGCTGAGAGACCTGCGCT 1849  
 Db 539 ATTGACTTAAGTGGATGGAGTGTGAGATCCTATGCAAGGCTGAGAGACCTGCGCT 598  
 QY 1850 TGAAGTCTG 1859  
 Db 599 TGAAGTCTG 608  
 RESULT 10  
 LOCUS BE334637 589 bp mRNA linear EST 14-JUL-2000  
 DEFINITION us1912.y1 NCI-CGAP\_Mam4 Mus musculus cDNA clone IMAGE:3257638 5' similar to TR:008854 008854 TRAF-INTERACTING PROTEIN ;, mRNA sequence.  
 ACCESSION BE334637  
 VERSION BE334637  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 589)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov/image/html/lresources.shtml](http://image.llnl.gov/image/html/lresources.shtml)  
 MGI:1069002  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 362.  
 Location/Qualifiers  
 1..589  
 FEATURES  
 source  
 /organism="Mus musculus"  
 /strain="NMRI"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3257638"  
 /clone\_id="NCI-CGAP\_Mam4"  
 /tissue\_type="tumor", gross tissue"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt; Site\_2: Not; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furch,

BASE COUNT 152 a 151 c 172 g 114 t  
 ORIGIN  
 Query Match 27.3%; Score 539; DB 10; Length 589;  
 Best Local Similarity 97.8%; Pred. No. 3e-101;  
 Matches 578; Conservative 0; Mismatches 10; Indels 3; Gaps 3;  
 NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996).  
 QY 681 ACAGTCACGGGTGAGACAGCTGCTGTACTGCTGCTTCCCTCAAGAAAGTATGAGAA 740  
 Db 1 ACAGTCACGGGTGAGACAGCTGCTGTACTGCTGCTTCCCTCAAGAAAGTATGAGAA 60  
 QY 741 TCTGAGAGACCTGGAGAGGCACAGGGAACTGGCTGACAGTGGTGAAGAGATTGGT 800  
 Db 61 TCTGAGAGACCTGGAGAGGCACAGGGAACTGGCTGACAGTGGTGAAGAGATTGGT 120  
 QY 801 GTCCCTAGAGACCAAGTTGAAGACTCTACACTGAGCTGGATCAGGCCAAGTTAGAACT 860  
 Db 121 GTCCCTAGAGACCAAGTTGAAGACTCTACACTGAGCTGGATCAGGCCAAGTTAGAACT 180  
 QY 861 GAGTCACGCCGAGAAGCACTTACAAGTGTCTGACCAAGATCAACAGCCTTAAGAAAGA 920  
 Db 181 GAGTCACGCCGAGAAGCACTTACAAGTGTCTGACCAAGATCAACAGCCTTAAGAAAGA 240  
 QY 921 GTCTGATGATCTCCAGGACCTTGAAGCCCTGCTCC-GGACCAATGAGAGGTGACGCC 979  
 Db 241 G-CTGATGATCTCCAGGACCTTGAAGCCCTGCTCCGGACCAATGAGAGGTGACGCC 299  
 QY 980 GCCTGCTTTTGAAGAGCCACCCCTGTGAGATGATGAACCCGAGCTTACCAAGCCAC 1039  
 Db 300 GCCTGCTTTTGAAGAGCCACCCCTGTGAGATGATGAACCCGAGCTTACCAAGCCAC 359  
 QY 1040 CCTTCGATGATGATGATGATCAATACACACTTTGATGAATACCCCTCAACCCAGA 1099  
 Db 360 CCTTCGATGATGATGATGATCAATACACACTTTGATGAATACCCCTCAACCCAGA 419  
 QY 1100 CCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1159  
 Db 420 CCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 479  
 QY 1160 TGCAGAAATGCTCCAGAGAGGTGACCAAGTCTCCAAGCCGAGTCCAGCTCACTGG 1219  
 Db 480 TGCAGAAATGCTCCAGAGAGGTGACCAAGTCTCCAAGCCGAGTCCAGCTCACTGG 539  
 QY 1220 GTGGCCAGCAATGTGTAGAGAGCTAGATGAGAACTGCTGCTGCTTC 1270  
 Db 540 GTGGCCAGCAATGTGTAGAGAGCTAGATGAGAACTGCTGCTGCTTC 589  
 RESULT 11  
 LOCUS AA684194 554 bp mRNA linear EST 09-DEC-1997  
 DEFINITION vm68410.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1003411 5' similar to TR:008854 008854 MTRIP. ;, mRNA sequence.  
 ACCESSION AA684194  
 VERSION AA684194  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 554)  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 554)  
 Maria, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.  
 TITLE The WashU-HMI Mouse EST Project  
 JOURNAL Unpublished (1996)  
 COMMENT Contact: Maria M/Mouse EST Project

WashU-HMI Mouse EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:567627  
Possible reversed clone: similarity on wrong strand  
High quality sequence stop: 512.  
Location/Qualifiers

## FEATURES

source

1..554  
/organism="Mus musculus"  
/strain="B6D2 F1/J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1003411"  
/tissue\_type="embryo"  
/dev\_stage="2-cell"  
/lab\_host="DH10B"  
/note="Organ: embryo; Vector: pBluescribe (modified);  
Site\_1: MluI; Site\_2: SalI; Cloned unidirectionally from  
mRNA prepared from 13,500 2-cell stage embryos. Primer:  
SalI(df): 5'-CGGTGACCGTCGACCGCTTTTCTTTT-3'.  
CDNAs were cloned into the MluI/SalI sites of a modified  
pBluescribe vector using commercial linkers (NEB).  
Average insert size: 1.2 kb."

BASE COUNT 149 a 151 c 139 g 115 t  
ORIGIN

Query Match 26.9%; Score 532.2; DB 9; Length 554;  
Best Local Similarity 99.3%; Pred. No. 7.8e-100;

Matches 545; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 53 CCACCAACTGTCCTGTCGTGCGAGCTGGTCCCTGGCTGCTTGAGTGCAGCATCA 112  
Db 6 CCACCAACTGTCCTGTCGTGCGAGCTGGTCCCTGGCTGCTTGAGTGCAGCATCA 65  
QY 113 TGCCTATCCTCTCTGTGACATCTGCTCCGACTTTTGATCATTCCCGTGCATGG 172  
Db 66 TGCCTATCCTCTCTGTGACATCTGCTCCGACTTTTGATCATTCCCGTGCATGG 125  
QY 173 CTGGCATCCAGTGGGCGACACTTTTCATGTCAGTATGCTAATCCAGTGGTTGAGACAG 232  
Db 126 CTGGCATCCAGTGGGCGACACTTTTCATGTCAGTATGCTAATCCAGTGGTTGAGACAG 185  
QY 233 CACCAAGTCGAGCTGCCACAGTGTAGATCCAGTGGGCAAAAGACTATTATAACA 292  
Db 186 CACCAAGTCGAGCTGCCACAGTGTAGATCCAGTGGGCAAAAGACTATTATAACA 245  
QY 233 AACCTTTTCTTGACCTGCCCGACGAGAGATGTCTTGATGACAGATTTCTTAAGA 352  
Db 246 AACCTTTTCTTGACCTGCCCGACGAGAGATGTCTTGATGACAGATTTCTTAAGA 305  
QY 353 ATGAAGTGAAGCGCTCAAGCTCAGCTTCCAGAAAGACAGGAGAAAGCGGACAGCC 412  
Db 306 ATGAAGTGAAGCGCTCAAGCTCAGCTTCCAGAAAGACAGGAGAAAGCGGACAGCC 365  
QY 413 AGGCATTATGACACTCTAGGGGACACCTGGAAGAACGCAATGCTACCTGAGTCC 472  
Db 366 AGGCATTATGACACTCTAGGGGACACCTGGAAGAACGCAATGCTACCTGAGTCC 425  
QY 473 TACAGAAAGCTTAAACAAGCAGAGATGCTGTTCACACCTT-6AAAAACAGTGAAG 531  
Db 426 TACAGAAAGCTTAAACAAGCAGAGATGCTGTTCACACCTTGAAGAAAAACAGTGAAG 485  
QY 532 TTCTGAGAGAGCGGACAGATGAGACAAAGCTGGGAGAGAGCCACCGACTCAG 591  
Db 486 TTCTGAGAGAGCGGACAGATGAGACAAAGCTGGGAGAGAGCCACCGACTCAG 545  
QY 592 TGCAAGATG 600

Db 546 TGCAAGATG 554  
|||||||

## RESULT 12

BG764118

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

702 bp mRNA linear EST 15-MAY-2001  
602737137F1 NIH\_MGC\_49 Homo sapiens CDNA clone IMAGE:4862312 5',  
mRNA sequence.  
BG764118  
BG764118.1 GI:14074771  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 702)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@rs-remail.nih.gov  
Tissue Procurement: ATCC/DCPD/DRP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM1722 row: 3 column: 09  
High quality sequence stop: 697.  
Location/Qualifiers

## FEATURES

source

1..702  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4862312"  
/tissue\_type="NIH\_MGC\_49"  
/tissue\_type="melanotic melanoma, high MDR (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pORF1; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Size-selected >500bp for average insert size  
1.8kb. Library constructed by Ling Hong in the laboratory  
of Gerald M. Rubin (University of California, Berkeley)  
using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
II RT (Life Technologies). Note: this is a NIH\_MGC  
Library."

BASE COUNT 183 a 180 c 202 g 136 t 1 others  
ORIGIN

Query Match 26.0%; Score 514.4; DB 10; Length 702;  
Best Local Similarity 87.4%; Pred. No. 3.6e-96;

Matches 574; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY 81 TGGTCCCTGGGCTGCTGAGTCGAGCATATGCTTCTCTCTGTCGACATATG 140  
Db 42 TGGTCCCTGGGCTGCTGAGTCGAGCATATGCTTCTCTCTGTCGACATATG 101  
QY 141 CTCCGACTTCTTCGATCACTCCCGTGCAGTGCATCCACTGTGGCCACACTTTCA 200  
Db 102 CTCCGACTTCTTCGATCACTCCCGTGCAGTGCATCCACTGTGGCCACACTTTCA 161  
QY 201 TCTGCAATGCTTAATCCAGTGTGTTGAGACAGCAGCAAGTGGAGCTGCCACAGTGA 260  
Db 162 CTTCGAGTGCCTTAATTCAGTGTGTTGAGACAGCAAGTGGAGCTGCCACAGTGC 221  
QY 261 AATCAGGTTGGCAAAAGACATTAATTAACAACCTTTCTTGACCTCGGCCAGGAGA 320  
Db 222 AATCAGGTTGGCAAAAGACATTAATTAACAACCTTTCTTGACCTCGGCCAGGAGA 281  
QY 321 GGAGAAATGCTTGGATGCAATTTCTTAAGAATGAAGTGAAGCGTCAAGCTCAGCT 380

|||||  
Db 282 GGAGATGCTTGGATGCGAGATTTCTTAAGATGACAGCAATGTCAGAGCCAGCT 341  
Oy 381 TTCCAGAAAGCAGGAGAAACGGAGACGAGCCATTATGACACTCTACGGAGAC 440  
Db 342 TTCCAGAAAGCAGGAGAAACGGAGACGAGCCATTATGACACTCTACGGAGATC 401  
Oy 441 CCTGGAGAAAGCAGGAGAAACGGAGACGAGCCATTATGACACTCTACGGAGAT 500  
Db 442 GTTGAAGAGAGCAGGAGAAACGGAGACGAGCCATTATGACACTCTACGGAGAT 461  
Oy 501 GCTGTGTTCCAGCTTGAAGAAACAGATGATGATCTCTGAGACAGCGCCAGATGAGACCA 560  
Db 462 GCTGTGTTCCAGCTTGAAGAAACAGATGATGATCTCTGAGACAGCGCCAGATGAGACCA 521  
Oy 561 ACAAGCTGGGAGGAGGAGCCAGCTCAAGTGAAGTGAAGAAACATGAGACCAATGA 620  
Db 522 ACAAGCACAAGAGAGAGGAGCCAGCTCAAGTGAAGTGAAGAAACATGAGACCAATGA 581  
Oy 621 GCTCTACTCCAGAGCCAGCTTCTGAGGTGAGGA-GATGATTCGAGACATGGGTGG 679  
Db 582 GCTTCTACTCCAGAGCCAGCTTCTGAGGTGAGGA-GATGATTCGAGACATGGGTGG 641  
Oy 680 GACACTCAGCGGTGAGAGAGCTGGCTGTGTTACTGCTGCTCCCTCAAGAAAGATG 736  
Db 642 GACAGACAGGGGTGAGAGAGCTGGGTGTGTACAGTGTGTCTCAAGAAAGATGACG 698

RESULT 13  
Bg682548 752 bp mRNA linear EST 01-MAY-2001  
LOCUS Bg682548 602624362P1 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4749447 5',  
DEFINITION mRNA sequence.  
ACCESSION Bg682548  
VERSION Bg682548.1 GI:13913945  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 752)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-rt@mail.nih.gov  
Tissue Procurement: James Cleaver, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

Plate: LMA10602 row: k column: 16  
High quality sequence start: 19  
High quality sequence stop: 736.  
Location/Qualifiers  
1..752

FEATURES  
source  
1..752  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4749447"  
/clone\_1id="NCI\_CGAP\_Skn4"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: skin; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.5kb. Library constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 200 a 194 c 212 g 146 t  
ORIGIN

Query Match

26.0%; Score 513.6; DB 10; Length 752;

Best Local Similarity 84.5%; Pred. No. 5.2e-96;  
Matches 612; Conservative 0; Mismatches 109; Indels 3; Gaps 3;

Oy 52 TCCACCAACATGTGTCTGTCTGTGACAGCTGTTCCCTGGGCTGTTAGTGCAGCCATC 111  
Db 27 TCTACGAAGCCGAGCTGTAGCATTTCTTGGCTGCTGGCCCTAGTCCACCATTC 86  
Oy 112 ATGCTTATCTCTCTCTCTGTCGACATATGTCGCCACTTCTTGATGATCCTCCGTCAGCTG 171  
Db 87 ATGCTTATCTCTCTCTGTCGACATATGTCGCCACTTCTTGATGATCCTCCGTCAGCTG 146  
Oy 172 GCTGCCATTCACATGTGGCCACACTTTTTCATCTGCAATGCTCAATTCAGATGTTTGAGACA 231  
Db 147 GCCGCATTCACATGTGGCCACACTTTTTCATCTGCAATGCTCAATTCAGATGTTTGAGACA 206  
Oy 232 GCACCAATCGGAGCTGCGCCACAGTGTAGATTCAGG-TTGCGCAAAAAGACTATTATATA 290  
Db 207 GCACCAATCGGAGCTGCGCCACAGTGTAGATTCAGG-TTGCGCAAAAAGACTATTATATA 266  
Oy 291 CAACTTTCTTTTGTGACTTCGCCACAGAGAGAGAGATGCTTGTGATGAGAAATTTTAA 350  
Db 267 TAAGCTCTTCTTTGATCTTGGCCAGAGAGAGAGATGCTTGTGATGAGAAATTTTAA 326  
Oy 351 GAATGAATGAGACAGCTCAAAAGCTCAAGCTTTCCAGAAAGACAGAGAAACGGAGAC 410  
Db 327 GAATGAATGAGACAGCTCAAAAGCTCAAGCTTTCCAGAAAGACAGAGAAACGGAGAC 386  
Oy 411 CCAGGCCATTTATCGACACTTACGGGACACCCCTGGAAGAAAGCAATGCTACCGTGAATC 470  
Db 387 CCAGGCCATTTATCGACACTTACGGGATGACCTGGAAGAAAGCAATGCTTGTGATG 446  
Oy 471 CTTACAGAAAGCCCTTAAACAGAGAGAGATGCTGTGTTCCACCTGGAAGAAAGAGATGA 530  
Db 447 TCTGACAGAGGCTTGGGCAAGGCCGAGATGCTGTGCTCCACTGGAAGAAAGAGATGA 506  
Oy 531 GTTCTGAGAGCGGAGATGAGACCAAAACAGCTGGAGAGAGGCCACGACCTCAA 590  
Db 507 GTACTTAGAGCAGCAGCAGATGAGACCAAAACAGCAGAGAGAGGCCCGGCTCAG 566  
Oy 591 GTGCAAGATGAAACCATGAGCAATGAGACTCTACTCCAGAGCCAGCTTGTGAGAT 650  
Db 567 GAGCAAGATGAAACCATGAGCAATGAGACTCTACTCCAGAGCCAGCTTGTGAGAT 626  
Oy 651 GGAGAGATGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 709  
Db 627 GGAGAGATGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 686  
Oy 710 ACTGCTGTGCTT-CAAGAAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 768  
Db 687 ACTGCTGTGCTT-CAAGAAAGATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 746  
Oy 769 GAAC 772  
Db 747 GCAC 750

RESULT 14  
BE304189 519 bp mRNA linear EST 13-JUL-2000  
LOCUS BE304189 601086421P1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:3500621 5',  
DEFINITION mRNA sequence.  
ACCESSION BE304189  
VERSION BE304189.1 GI:9174333  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 519)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHM8559 row: 1 column: 06  
 High quality sequence stop: 515.

## FEATURES

## source

Location/Qualifiers  
 1. 519

/organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3500621"  
 /clone\_lib="NCI\_CGAP\_Mam6"  
 /sex="Female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Jeffrey Green, M.D., NIH"

BASE COUNT 159 a 116 c 152 g 92 t  
 ORIGIN

Query Match 26.0%; Score 513.4; DB 10; Length 519;  
 Best Local Similarity 99.8%; Pred. No. 6e-96;

Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 263 TCCAGGTTGGCAAAAGCTATTAACAACATTTTGTGACCTGCCCGAGAGAGG 322  
 |||||||  
 Db 5 TCCAGGTTGGCAAAAGCTATTAACAACATTTTGTGACCTGCCCGAGAGAGG 64  
 |||||||  
 QY 323 AGAATGCTTGGATGACAGATTTCTTAAGATGAAGTGGACGGCTCAAGCTCAGCTT 382  
 |||||||  
 Db 65 AGAATGCTTGGATGACAGATTTCTTAAGATGAAGTGGACGGCTCAGCTT 124  
 |||||||  
 QY 383 CCCGAAAGACAG 442  
 |||||||  
 Db 125 CCCGAAAGACAG 184  
 |||||||  
 QY 443 TGGAGAAAGCAATGCTACCGTGGAGTCCCTACAGAGAGAGAGAGAGAGAGAG 502  
 |||||||  
 Db 185 TGGAGAAAGCAATGCTACCGTGGAGTCCCTACAGAGAGAGAGAGAGAGAGAG 244  
 |||||||  
 QY 503 TGTGTTCCACCTGAAAAAAGATGAAGTCTGTGAGAGAGAGAGAGAGAGAGAG 562  
 |||||||  
 Db 245 TGTGTTCCACCTGAAAAAAGATGAAGTCTGTGAGAGAGAGAGAGAGAGAGAG 304  
 |||||||  
 QY 563 AAGCTCGGAG 622  
 |||||||  
 Db 305 AAGCTCGGAG 364  
 |||||||  
 QY 623 TCCCTACTCAAG 682  
 |||||||  
 Db 365 TCCCTACTCAAG 424  
 |||||||  
 QY 683 AGTCAGCGGTGGAG 742  
 |||||||  
 Db 425 AGTCAGCGGTGGAG 484  
 |||||||  
 QY 743 TGAAGGAAGCTCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777  
 |||||||  
 Db 485 TGAAGGAAGCTCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519  
 |||||||

RESULT 15  
 BGI20736 BGI20736 784 bp mRNA linear EST 30-JAN-2001  
 LOCUS

DEFINITION 602347389F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4442038 5',  
 mRNA sequence.  
 ACCESSION BGI20736  
 VERSION BGI20736.1 GI:12614245  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 784)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apbs-remail.nih.gov

Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHM10214 row: b column: 23  
 High quality sequence stop: 713.

## FEATURES

Location/Qualifiers  
 1. 784

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4442038"  
 /clone\_lib="NIH\_MGC\_90"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; Oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 207 a 200 c 224 g 153 t  
 ORIGIN

Query Match 25.9%; Score 510.6; DB 10; Length 784;  
 Best Local Similarity 87.2%; Pred. No. 2.2e-95;

Matches 594; Conservative 0; Mismatches 84; Indels 3; Gaps 3;

QY 81 TGTGTTCCCTGGGCTGCTTGAAGTCAATGCTATCTCTCTGTGCACTATCTG 140  
 |||||||  
 Db 62 TGTGTTCCCTGGGCTGCTTGAAGTCAATGCTATCTCTCTGTGCACTATCTG 121  
 |||||||  
 QY 141 CTCGACCTTCTTGATCACTCCCGTGAGCTGGCTGCATTCAGTGGCTGCTGCTG 200  
 |||||||  
 Db 122 CTCGACCTTCTTGATCACTCCCGTGAGCTGGCTGCATTCAGTGGCTGCTGCTG 181  
 |||||||  
 QY 201 TCTGCATATGCTTATTCAGTGTGAGACAGACCAAGTGGAGCTGCCACAGTGTAG 260  
 |||||||  
 Db 182 TCTGCATATGCTTATTCAGTGTGAGACAGACCAAGTGGAGCTGCCACAGTGTAG 241  
 |||||||  
 QY 261 AATCCAGTTGGCAAAAGACTATTATTAACAACCTTTTCTTGAAGTCTGCCCCAGAGA 320  
 |||||||  
 Db 242 AATCCAGTTGGCAAAAGACTATTATTAACAACCTTTTCTTGAAGTCTGCCCCAGAGA 301  
 |||||||  
 QY 321 GGAGAAATCTTGGATGAGATTTCTTAAGAAATGAAGTGGACAGAGCAAAAGTCCAGCT 380  
 |||||||  
 Db 302 GGAGAAATCTTGGATGAGATTTCTTAAGAAATGAAGTGGACAGAGCAAAAGTCCAGCT 361  
 |||||||  
 QY 381 TTCCGAGAAAGACAG 440  
 |||||||  
 Db 362 TTCCGAGAAAGACAG 421  
 |||||||  
 QY 441 CCTGGAAGAGCAATGCTACCGTGGAGTCCCTACAGAGAGAGAGAGAGAGAGAG 500  
 |||||||  
 Db 422 GCTGGAAGAGCAATGCTACCGTGGAGTCCCTACAGAGAGAGAGAGAGAGAGAGAG 481  
 |||||||

Oy	501	GCGTGTTCACCCCTGAAAAAACAGATGAATTCCTGGACACCAGCGCAGGATATACACCAA	560
Db	482	GCTGTGCTCCCACTGAAAAAGCATGAAGTAATTGAGCACACAGCGGATTAGACCAA	541
Oy	551	ACAAGCTCGGGAGGAGGCCACCACACTCAAGTGCAAGATATAAAACCATGAGCAAAATTGA	620
Db	542	ACAAGACACAAAGAGGAGGCC - GCGGCTCAGSAGACAGATGAAGACCATGAGACGATTTGA	600
Oy	621	GCTCTACTCCAGAGCCACGGTTCTGAGTGTGGAGAGATATATTCGAGACATGGGTGTGG	680
Db	601	GCTTCTACTCCAGAGCCACGGCCCTTGAGTGTGGAGAGATATATCCGAGACATGGGTGTGG	660
Oy	681	ACAGTTCAGCGGTGGAGCAGCTGTGCTGTACTCGGTCCCTCCATAAAGAAAGATATAGAA	740
Db	661	ACAGTTCAGCGGTGGAGACCGGCTGG - TGTGTACTGTGTGTCTCTCCAGAAGAGT - CCAAAA	718
Oy	741	TCTGAGGAAGCTCGGAAGGC	761
Db	719	TCTTAAAGAAAGCCGGGAGGCC	739

RESULT	16
BG420765	
LOCUS	BG420765
DEFINITION	908 bp mRNA linear EST 14-MAR-2001 602445538t1 NIH_MGC_14 Homo sapiens CDNA clone IMAGE:4586846 5'
VERSION	mRNA sequence.
ACCESSION	BG420765
KEYWORDS	BG420765.1 GI:13327271
SOURCE	EST. human.

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 908)
TITLE	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
Plate: LNCM316 row: d column: 15  
High quality sequence stop: 719.

FEATURES	Location/Qualifiers
source	1. .908

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:4586846"
/clone_1ib="NH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: POT7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

Query Match	25.5%	Score 504.2;	DB 10;	Length 908;
Best Local Similarly	82.0%	Pred. No. 4.4e-94;		
Matches 663; Conservative	0;	Mismatches 93;	Indels 53;	Gaps 5;

91 GGCTGCTGAGTCGAGCCATCATGCCCTATCCTCTCTCTGTGTCACCTATCTGCTCCGACTTC 150

[illegible]

RESULT	17			
LOCUS	BE546959	825 bp	mrna	linear
DEFINITION	60107192991 NIH_MGC_12 Homo sapiens CDNA clone IMAGE:3458023 5',			
ACCESSION	BE546959			
VERSION	BE546959.1 GI:9775604			
KEYWORDS	EST.			
SOURCE	human.			

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
 Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.  
 1 (bases 1 to 825)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.

Email: c9apbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LAM8448 row: j column: 08

High quality sequence stop: 655.

## FEATURES

## Source

Location/Qualifiers

1..825

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3458023"

/clone\_1lb="NIH\_MGC\_12"

/tissue\_type="cervical carcinoma cell line"

/lab\_host="DH10B"

/note="Organ: cervix; Vector: PCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

## BASE COUNT

230 a 209 c 242 g 144 t

## ORIGIN

## Query Match

25.5%; Score 503.2; DB 10; Length 825;

Best Local Similarity 85.1%; Pred. No. 7.2e-94;

Matches 633; Conservative 0; Mismatches 103; Indels 8; Gaps 6;

## FEATURES

## Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

## BASE COUNT

167 a 174 c 177 g 127 t

## ORIGIN

## Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

## FEATURES

## Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DJ012YH21"

/clone\_1lb="LTI\_NFL008\_Tc2"

/sex="male"

/tissue\_type="T cells from T cell leukemia"

/note="Vector: pcMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pcMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact: Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA fax: (1) 301 610 8371

Email: [liang@lifetech.com](mailto:liang@lifetech.com) URL: <http://fulllength.invitrogen.com>

BASE COUNT

167 a 174 c 177 g 127 t

ORIGIN

Query Match

25.2%; Score 497.4; DB 9; Length 645;

Best Local Similarity 88.4%; Pred. No. 1.2e-92;

Matches 540; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

FEATURES

Source

Location/Qualifiers

1..645



|||||  
Db 275 GGAGATGCTGTGATGCAAGATCTCTTAAGATGAACATGACAAATGTCAGAGCCAGCT 334  
Oy 381 TTCCCGAAGAACAGAGAGAAACGGAGACAGCCAGCCATTATCGACACTCTACGGGAC 440  
Db 335 TTCCCGAAGAACAGAGAGAAACGGAGACAGCCAGCCATTATCGACACTCTACGGGATC 394  
Oy 441 CCTGGAAGAACGCAATGCTACCGTGGAGTCCCTACAGAACGCTTAAACAGAGAGAT 500  
Db 395 GCTGGAAGAACGCAATGCTACCTGTGTGTATCTGTGACAGCAGGCTTGGGCAAGCCGAGAT 454  
Oy 501 GCTGTGTTCCACCTGAAAAAACAGATGAAGTTCTTGAGACAGCGCAGATGAGACCA 560  
Db 455 GCTGTGTTCCACCTGAAAAAACAGATGAAGTTAGAGCAGCAGAGATGAGACCA 514  
Oy 561 ACAACCTGGGAGAGAGCCACACCGACTCAAGTGAAGATGAACACCATGAGCAATTTGA 620  
Db 515 ACAACCAAG 574  
Oy 621 GCTCTACTCCAGAGCCAGCGCTTGTGAGTGGAGAGAGATGATTCGAGACATGGTGG 680  
Db 575 GCTTCTACTCCAGAGCCAGCGCGCTGAGAGTGGAGAGAGATGCCAGACATGGTGG 634  
Oy 681 ACAGTCACGG 691  
Db 635 ACAGTCACGG 645

RESULT 19  
BE869186 888 bp mRNA linear EST 20-OCT-2000  
LOCUS 601445180P1 NIH\_MGC\_65 Homo sapiens cDNA clone IMAGE:3849451 5',  
DEFINITION mRNA sequence.  
ACCESSION BE869186  
VERSION BE869186.1 GI:10317962  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 888)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsab@remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.nih.gov  
Plate: LHAM9566 row: O column: 20  
High quality sequence stop: 692.  
Location/Qualifiers  
1..888  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3849451"  
/clone\_id="NIH\_MGC\_65"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.8 kb. Library constructed by Life  
Technologies."

BASE COUNT 254 a 199 c 268 g 167 t  
ORIGIN

Query Match 24.8%; Score 490.6; DB 10; Length 888;  
Best Local Similarity 79.8%; Pred. No. 2.9e-91;

Matches 665; Conservative 0; Mismatches 159; Indels 9; Gaps 7;  
Oy 319 GAGGAGATGCTGTGATGCAAGATTTCTTAA- GAATGAATCGACACAGCTCAAGCTCA 377  
Db 8 GAGGAGATGCTGTGATGCAAGATTTCTTAAAGATGAATCACTGAGCAATGTGAGGCCA 67  
Oy 378 GCTTTCAGAAAGACAGAGAGAAACGGAGACAGCCAGCCATTATCGACACTCTACGGGA 437  
Db 68 GCTTTCAGAAAGACAGAGAGAAACGGAGACAGCCAGCTCATCATGACACTCTCGGGGA 127  
Oy 438 CACCTGGAAGAGAGC- AATGCTACCGTGAAGTCCCTACAGAACGCTTAAACAAGGAC 496  
Db 128 TACGCTGGAAGAGAGCAGATGCTACTGTGTATCTCTCAGAGGCTTGGGCAAGGCG 187  
Oy 497 AGATGCTGTGTTCCACCCCTGAAAAACAGATGAAGTTCTCGAGAGCGGAGAGATGGA 556  
Db 188 AGATGCTGTGTTCCACACTGAAAAAGCAGATGAAGTACTTGAAGCAGCAGAGATGGA 247  
Oy 557 CCNAAAGAGCTCGGAGAGAGGCCACCGACTCAAGTGAAGATGAACCAATGAGAGCAA 616  
Db 248 CCNAAAGAGCAG 307  
Oy 617 TTGAGCTCTTACTCCAGAGCCAGCGCTTGAAGTGGAGAGATGATTCGAGACATGGTG 676  
Db 308 TTGAGCTCTTACTCCAGAGCCAGCGCCTGAGTGGAGAGATGATTCGAGACATGGTG 367  
Oy 677 TGGGACACTCAGCGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 736  
Db 368 TGGGACACTCAGCGGTGGAG 427  
Oy 737 AGAATCTGAAG 796  
Db 428 AGAATCTGAAG -T 486  
Oy 797 TGTGTCTCTTAGAGAGCAAGTTGAAGACTCTCAACACTGAGCTGATGAGCCAAATTAG 856  
Db 487 TGTGTCTCTTAGAGAGCAAGTTGAAGACTCTCAACACTGAGCTGATGAGCCAAATTAG 546  
Oy 857 AACTGAGTCAAGCCAG 916  
Db 547 AACTGAGTCAAGCCAG 606  
Oy 917 AGAATCTGAAG 973  
Db 607 AGAATCTGAAG 666  
Oy 974 TCAGCCGCTGTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032  
Db 667 TCAGCCGCTGTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726  
Oy 1033 CAGCCAGCCCTTCGGATGAGAGAT- TGATCTCAATACCACTTGTGATGTAATACCCCTCC 1091  
Db 727 GGGCATCTATCCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785  
Oy 1092 AACCCAGAGCTCTGTGCTCCAGAGATTCCTCCCAAGAGAGAGAGAGAGAGAGAGAG 1144  
Db 786 CAGCCGCGCTCAGAGTTCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 838

RESULT 20  
B1758046 668 bp mRNA linear EST 25-SEP-2001  
LOCUS 603030577P1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5200986 5',  
DEFINITION mRNA sequence.  
ACCESSION B1758046  
VERSION B1758046.1 GI:15749624  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 668)  
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.



Db 303 GGAGATCTCTTGATGCAAGATTCTTAAGATGAAGTGGACAATGTGACAGCCCACT 362  
QY 381 TTCCAGAAAGACAGAGAAACGGAGACGCCATTAATGACACTTACGGACAC 440  
Db 383 TTCCAGAAAGACAGAGAAACGGAGACGCCATTAATGACACTTACGGAGATAC 422  
QY 441 CCTGAGACAGCAATGCTTACCGTGGAGTCCCTACAGAACGCCGAGAGAT 500  
Db 423 GCTGAAACAGCAATGCTTACCGTGGAGTCCCTACAGAACGCCGAGAT 482  
QY 501 GCTGTTCCACCCGAAAGAAAGATGAAGTCTCTGAGACAGCGGACGATGAGACCA 560  
Db 483 GCTGCTCAGACAGTGAAGAAAGACAGATGAAGTGAAGCAAGATGAGACCA 542  
QY 561 ACAAGCTGGAGAGAGGCCACCGACACTCAAGTGAAGTGAAGCAAGATGGA 620  
Db 543 ACAAGCAACAGAGAGGCCCGCGCTCAGAGCAAGATGAAGCAAGATGGA 602  
QY 621 GCTCTACTCCAGACCGCTTCTGAGGTGAGAGATGATTCGAG-ACATGGGTGG 679  
Db 603 GCTTCTACTCCAGAGCCAGCG-CCGTGAGGTGAGAGATGATTCGAGAACATGGGTGG 661  
QY 680 GACAGTCAGCGGTGAGAGAGTGGCTGTGT-ACCTGCTGTCCCTCAAGAAAGATGAG 738  
Db 662 GACAGTCAGCGGGGAACCGTGGCTGTGTAAAGGGGGTCTCCAGAGAAAGCTCCAA 721  
QY 739 AATCTGAAGAGCTCGGAGGCCA 763  
Db 722 AATCCAAACAGAGGCCGAGAGCTCA 746

RESULT 22  
BF219707 632 bp mRNA linear EST 08-NOV-2000  
LOCUS 601296627F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:282107 5',  
DEFINITION mRNA sequence.

ACCESSION BF219707  
VERSION BF219707.1 GI:11125801  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 632)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabps-remail.nih.gov](mailto:cgabps-remail.nih.gov)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: IRAL1 row: e column: 07  
High quality sequence stop: 630.  
Location/Qualifiers  
1. 632

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:282107"  
/clone\_lib="NIH\_MGC\_7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adapter: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 163 a 171 c 172 g 126 t  
ORIGIN

Query Match 24.1%; Score 476.2; DB 10; Length 632;  
Best Local Similarity 87.7%; Pred. No. 2.8e-88;  
Matches 520; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 81 TGGTCCCTGGGGCGCTTGAAGTCCGACATCATGCTTCTCTGTCCTATCTG 140  
Db 37 TGGCTGCTGGGGCCCTTGAAGTCCGACATCATGCTTCTCTGTCCTATCTG 96  
QY 141 CTCGACTTCTTCGATCACTCCCGTGAAGTGGTCCATCCATCGTGGCCACATTTTCA 200  
Db 97 CTCGAGCTTCTTCGATCACTCCCGTGAAGTGGTCCATCCATCGTGGCCACATTTTCA 156  
QY 201 TCTGCAATGCTTAATTCAGTGGTTTGAGACAGCACAAGTGGACCTGCCACAGTTAG 260  
Db 157 CTTCAGTGCCTAATTGATGCTTGTGAGACAGCACAAGTGGACCTGCCACAGTTAG 216  
QY 261 AATCCAGGTGGCAAAAGACTATTATTAACAACCTTTCTTGACCTGCCGAGAGA 320  
Db 217 AATCCAGGTGGCAAAAGACTATTATTAACAACCTTTCTTGATCTTGCCAGAGAGA 276  
QY 321 GGAGAATGCTTTGGATGCGAGATTTCTTAAGAAATGAAGTGGACAGCTCAAGCTCAGCT 380  
Db 277 GGAGAATGCTTTGGATGCGAGATTTCTTAAGAAATGAAGTGGACAGCTCAGCT 336  
QY 381 TTCCAGAAAGACAGAGGAAACGGAGACAGCGCATTTATGCACTACGGGACAC 440  
Db 337 TTCCAGAAAGACAGAGGAAACGGAGACAGCGCATTTATGCACTACGGGATAC 396  
QY 441 CCTGGAAGAACGCAATGCTACCGTGAAGTCCCTGACAGACGCTTAACAAAGGAGAT 500  
Db 397 GCTGGAAGAACGCAATGCTACCGTGAAGTCCCTGACAGACGCTTGGGCAAGCGAGAT 456  
QY 501 GCTGTGTTCCACCTGAAAGAAACAGATGAAGTCTCTGAGACAGCGGAGATGAGACCA 560  
Db 457 GCTGTGTTCCACCTGAAAGAAACAGATGAAGTCTCTGAGACAGCGGAGATGAGACCA 516  
QY 561 ACAAGCTGGAGAGAGGCCACCGACACTCAAGTGAAGTGAAGCAATGGAAGCAATTTGA 620  
Db 517 ACAAGCAACAGAGAGGCCCGCGCTCAGAGCAAGATGAAGCAATGGAAGCAATTTGA 576  
QY 621 GCTCTCTCCAGAGCGGCTTCTGAGGTGAGAGATGATTTGAGACATGG 673  
Db 577 GCTTCTACTCCAGAGCCAGCGCCTGAGGTGAGAGATGATTCGAGACTGGG 629

RESULT 23  
BG117979 912 bp mRNA linear EST 30-JAN-2001  
LOCUS 602351223F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:4446149 5',  
DEFINITION mRNA sequence.

ACCESSION BG117979  
VERSION BG117979.1 GI:12611485  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 912)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabps-remail.nih.gov](mailto:cgabps-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov  
Plate: L1AM10224 row: n column: 06  
High quality sequence stop: 678.  
Location/Qualifiers

## FEATURES

source

1. 912  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4446149"  
/clone\_lib="NIH.MGC.90"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Liver; Vector: pCMV-SPO0R6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH.MGC Library."

BASE COUNT 236 a 243 c 273 g 160 t  
ORIGIN

Query Match 23.3%; Score 459.8; DB: 10; Length 912;  
Best Local Similarity 83.8%; Pred. No. 6.0e-85;

Matches 591; Conservative 0; Mismatches 107; Indels 7; Gaps 6;

81 TGGTTCCCTGGGCTGTTGAGTCGAGCCATCATGCTATCTCTCTGTCACATCTG 140  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db TGGCTGCTGGGCCCCCTTGAATCCAGCCATCATGCTATCTCTCTGTCACATCTG 124  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 141 CTCGACCTCTTTCGATCAGTCCCGTGAAGTGGCCATCCATCGTGGCCACATTTTCA 200  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 125 CTCGACCTCTTTCGATCAGTCCCGTGAAGTGGCCATCCATCGTGGCCACATTTTCA 184  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 201 TCTGCATTCCTATTCAGTGGTTTGAGACAGCACAAGTGGAGCTGCCACAGTGTAG 260  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 185 CTTCAGTGCCTAATTAAGTGTGGAGACAGCACAAGTGGAGCTGCCACAGTGTAG 244  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 261 AATCCAGTGGCAAAAGAACTATTATTAACAACCTTTCTTTCAGCTGGCCAGAGA 320  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 245 AATCCAGTGGCAAAAGAACTATTATTAACAACCTTTCTTTCAGCTGGCCAGAGA 304  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 321 GGAAGATGCTTGGATGAGATTTCTTAAGATGAACTGACAGCGTCAAGCTCA-GC 379  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 305 GGAGAATGCTTGGATGAGATTTCTTAAGATGAACTGACAGCGTCAAGCTCAAGCTCA 364  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 380 TTTCGCAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 439  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 365 TTTCGCAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 424  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 440 CCTGGAAGAAAGCAATGCTACCGTGAAGTCCCTACAGAACGCTTAACAAGCA-GAG 498  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 425 CGGTGAAGAAAGCAATGCTACCGTGAAGTCCCTACAGAACGCTTAACAAGCA-GAG 484  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 499 ATGCTGTCTCCACCTGAAAAAAGATGAACTCTGAGACAGGCGGAGAGTGAAGC 558  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 485 ATGCTGTCTCCACCTGAAAAAAGATGAACTCTGAGACAGGCGGAGAGTGAAGC 544  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 559 AAAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 618  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 545 AAAAAGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 619 GAGCTCTCTACT-CCAGAGCCAGCGTTCTGAGGTGAGAGAG-ATGATTTCAGACAT- 675  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 605 GAGCTCTCTACTCCAGAGCCAGCGTTCTGAGGTGAGAGAG-ATGATTTCAGACAT- 664  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 676 GTGGAGAGTACAGCGGT--GGAGCAGCTGCTGTCTACTGCGTGTCTCTCAAGAAAGT 733  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 665 GTGGAGAGTACAGCGGTGGAGAGCAGCTGTGTCTACTGCGTGTCTCTCAAGAAAGT 724  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 734 ATGGAATCTGAAGGAACCTGGAAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 778  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 725 ACGACATCTACAGAGGCCCGGAGGCTCTCAGGAGAAAGTGGGTG 769  
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 24  
BB821649  
LOCUS

BB821649 528 bp mRNA linear EST 19-NOV-2001  
BB821649 RIKEN full-length enriched, mammary gland RCB-0526  
Jyg-MC(A) CDNA Mus musculus CDNA clone G830012P09 3, mRNA  
sequence.

## ACCESSION

VERSION BB821649  
KEYWORDS BB821649.1 GI:16994278

## SOURCE

ORGANISM

## REFERENCE

AUTHORS

Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,  
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,  
Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,  
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,  
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,  
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,  
Watanabe,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)

## JOURNAL

COMMENT

Unpublished (2001)  
Contact: Yoshinide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,  
M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wag1,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,  
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,  
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,  
Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for  
further details.  
e mouse tissues.

## FEATURES

source

1. 528  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="G830012P09"  
/clone\_lib="RIKEN full-length enriched, mammary gland  
RCB-0526 Jyg-MC(A) CDNA"  
/tissue\_type="mammary gland"  
/cell\_line="RCB-0526 Jyg-MC(A)"

## BASE COUNT

137 a 121 c 117 g 153 t

Query Match 22.1%; Score 437.2; DB: 9; Length 528;  
Best Local Similarity 93.2%; Pred. No. 3.3e-80;  
Matches 490; Conservative 0; Mismatches 33; Indels 3; Gaps 3;

QY 1434 TAACTCCAGGCCAAGATTAACAGAAAGTGAATTAAGACTGTGATCTGCTCCCA 1493  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2 TAAGTCCAGGCCAAGAAATTAACAGAAATTGAGAAATTAAGACTGTGAGTTCCTCCCA 61

Qy 1494 -GCCCAAGCTGGATACCTTCTTATGTAGTGAAGCGTGACCGAGTGAATGTTTGCATTTA 1552  
|||||

Db 62 GGGCCAACTGGATACCTTTTATTTTCACTGACCTGACCGAGATCATGTTTGCATTTA 121  
|||||

Qy 1553 GTGGGCCAAGACCTGGCTTAACCGGAGTGTGTTTGAAGATGGCTCTTTGGACAGTC 1612  
|||||

Db 122 GTGGGCCAAGACCTGGCTTAACCGGAGTGTGTTTGAAGATGGCTCTTTGGACAGTC 181  
|||||

Qy 1613 CAAGAGATGGCCAGAAACACACTTCTGTTCACTGGC -CCCTGACACACTGGG 1671  
|||||

Db 182 CAAGGAATATGCCAGAAACACACTTCTGTTCACTGGC -CCCTGACACACTGGG 241  
|||||

Qy 1672 AAGCCATGACCACTGTTACTGTTCCGATCAG -CAGGCGCTACTCCAGTTGACAGGTTT 1730  
|||||

Db 242 AAGCCATGACCACTGTTACTGTTCCGATCAG -CAGGCGCTACTCCAGTTGACAGGTTT 301  
|||||

Qy 1731 TGCCTATAGCTAACAACAGGTGCTGACTCTTTTGTGTTTATAGAACAGGTTACA 1790  
|||||

Db 302 TGCCTATAGCTAACAACAGGTGCTGACTCTTTTGTGTTTATAGAACAGGTTACA 361  
|||||

Qy 1791 TTGACTTAAGTGATGGAGTGGAGTGGAGATCCTATGACAGGCTGAGACCTGCGCTT 1850  
|||||

Db 362 TTACTGTAAGTGAATGGAGTGGAGTGGAGATCCTATGACAGGCTGAGACCTGCGCTT 421  
|||||

Qy 1851 GAACCTCTGCTGCTCCAGCTTATGCTTGAATATATGGGTGAGTGTATAGGAA 1910  
|||||

Db 422 AAACCTCTGCTGCTCCAGCTTATGCTTGAATATATGGGTGAGTGTATAGGAA 481  
|||||

Qy 1911 AGCTTGGGGAAGTTTCTGTGTAATAAATGAAGGATCTTTTCTTC 1956  
|||||

Db 482 AGCTTGGGGAAGTTTCTGTGTAATAAATGAAGGATCTTTTCTTC 527  
|||||

RESULT 25  
AM490717/c 441 bp mRNA linear EST 24-FEB-2000  
LOCUS  
DEFINITION  
UI-M-BH3-asx-c-07-0-UI.s1 NIH\_BMAP\_M.S4 Mus musculus cDNA clone  
AM490717  
UI-M-BH3-asx-c-07-0-UI 3', mRNA sequence.  
ACCESSION  
AM490717.1 GI:7060988  
VERSION  
AM490717.1  
KEYWORDS  
EST.  
SOURCE  
house mouse.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 441)  
AUTHORS  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
COMMENT  
Contact: Chln, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mstt@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to identify it as a clone from the  
normalized prefrontal cortex library cDNA library Preparation: M.B.  
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
clones from RESEARCH GENERICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMAP cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements  
Seq primer: M13 Forward  
POLYA=yes.

FEATURES  
source

Location/Qualifiers  
1..441  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH3-asx-c-07-0-UI"  
/clone\_id="NIH\_BMAP\_M.S4"  
/dev\_stage="27-32 days"  
/db\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site\_1: Not I; Site\_2: Eco RI; The  
NIH\_BMAP\_M.S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged  
normalized libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH\_BMAP\_M.S4,  
NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, NIH\_BMAP\_M.S3.1,  
NIH\_BMAP\_M.S2, NIH\_BMAP\_M.S1. The subtracted library  
(NIH\_BMAP\_M.S4) was constructed as follows: PCR amplified  
cDNA inserts from NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, and  
NIH\_BMAP\_M.S3.1 clones from which 3' ESTs had been derived  
was used as a driver in a hybridization with a pool of  
the NIH\_BMAP\_M.S3.3, NIH\_BMAP\_M.S3.2, and NIH\_BMAP\_M.S3.1  
libraries in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library)  
was purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (Life Technologies) to generate the  
NIH\_BMAP\_M.S4 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research  
6:791-806, 1996)  
TAG\_LIB=NIH\_BMAP\_M.S4  
TAG\_TISSUE=prefrontal.cortex  
TAG\_SEQ=CGTCA"

BASE COUNT 117 a 115 c 94 g 115 t  
ORIGIN

Query Match 21.9%; Score 433; DB 9; Length 441;  
Best Local Similarity 98.9%; Pred. No. 2.5e-79;  
Matches 436; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1534 AGAGTGATGTTTGGCAATTAATGAGGCGCAAGACCTGGCTAACCGGAAAGTTTTGAAGAT 1593  
|||||

Db 441 AGAGTGATGTTTGGCAATTAATGAGGCGCAAGACCTGGCTAACCGGAAAGTTTTGAAGAT 382  
|||||

Qy 1594 GGCCTGCTTGGAGCAGGCGCAAGAGAGTGGCCGAAACACACTTCTGTTGCTACATGC 1653  
|||||

Db 381 GGCCTGCTTGGAGCAGGCGCAAGAGAGTGGCCGAAACACACTTCTGTTGCTACATGC 322  
|||||

Qy 1654 GCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATCAGACAGGCGCTAC 1713  
|||||

Db 321 GCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATCAGACAGGCGCTAC 262  
|||||

Qy 1714 TTCAGTTGCAAGGTTTGTCTTATAGCTACACAGGTGAGCTGACACTCTTTTGT 1773  
|||||

Db 261 TTCAGTTGCAAGGTTTGTCTTATAGCTACACAGGTGAGCTGACACTCTTTTGT 202  
|||||

Qy 1774 TATGAACAGGCGTACATGACTCTAGTGGATGGAGTCTGAGAGATCCTATGACAGGC 1833  
|||||

Db 201 TATGAACAGGCGTACATGACTCTAGTGGATGGAGTCTGAGAGATCCTATGACAGGC 142  
|||||

Qy 1834 TGGAGACCTGCGCTTGAACCTCTGCTGCCTCCACACTTATGCTGAATTAATGGGGT 1893  
|||||

Db 141 TGGAGACCTGCGCTTGAACCTCTGCTGCCTCCACACTTATGCTGAATTAATGGGGT 82  
|||||

Qy 1894 GAGGTGCTGATAGGAAAGTGGGGAAGTTTCTGTGTAATAAATGAAGGATCTTTTC 1953  
|||||

Db 81 GAGGTGCTGATAGGAAAGCTGGGGAAGTTTCTGTGTAATAAATGAAGGATCTTTTC 22  
|||||







Db 459 GGAGGACGAACAAATTCATCCAGCTAGGAGACACCACTATTCCAGCAGCTCTT 400  
 QY 1435 AAGTCCAGGCGCAGAGTAAGAGAGAGTAAGAGTGTGAGTTCGCTCCAG 1494  
 Db 399 AAGTCCAGGCGCAGAGTAAGAGAGAGTAAGAGTGTGAGTTCGCTCCAG 340  
 QY 1495 CCCAAGCTGATACCTTCTATGTCACTGAGACGCTGACAGAGTATGTTGCAATTAGT 1554  
 Db 339 CCCAAGCTGATACCTTCTATGTCACTGAGACGCTGACAGAGTATGTTGCAATTAGT 280  
 QY 1555 GGGCAGAGCTGCTTCAACCGGAAGTCTTTTGGAGAAGTGTCTCTTGAGCAGTCCA 1614  
 Db 279 GGACCAAGCTGCTGCTCAACCGGAAGTCTTTTGGAGAAGTGTCTCTTGAGCAGTCCA 222  
 QY 1615 AGAGAGATGCCCAAAACACACTTCTGTGTACAGTGGCGCCGACACAGTGGGAG 1674  
 Db 223 -----TTCCGTGTGTACAGTGGCGCCGACACAGTGGGAG 187  
 QY 1675 CCACATGACAGCTTACTGTTCGATCAGCAGGCTACTTCCAGTTCAGAGGTTTGTCT 1734  
 Db 186 CCAGGTGACAGCTTACTGTTCGATCAGCAGGCTACTTCCAGTTCAGAGGTTTGTCT 127  
 QY 1735 TATAGTACAAACAGGTGTGCTGCTGCTCTTTTGTATAGAACAGGCTCACTTGA 1794  
 Db 126 TATAGTACAAACAGGTGTGCTGCTGCTCTTTTGTATAGAACAGGCTCACTTGA 67  
 QY 1795 CTCAAGTGTGAGAGTGGAGTGGAGAGTCCATGACAGCTGAGAGACCTGCTTGAAC 1854  
 Db 66 CTCTCAGTGTGAGAGTGGAGAGTGGAGAGCCTATGACAGCTGGA-GACCTGCTTGAAC 8

RESULT 29  
 BE031265 547 bp mRNA linear EST 09-JUL-2000  
 LOCUS  
 DEFINITION 129770 MARRC 1P16 Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BE031265  
 VERSION BE031265.1 GI:8326274  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 1 (bases 1 to 547)  
 Fahrenkrug,S.C., Feking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
 Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.  
 and Keele,J.W.  
 Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 Unpublished (2000)  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACACGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCAGCAGC  
 Plate: 65 row: H column: 4  
 Seq primer: ATTAGGTGACACATATAG.

FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..347  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARRC 1P16"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
 library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."

BASE COUNT 166 a 130 c 161 g 90 t  
 ORIGIN

Query Match 21.5%; Score 424.4; DB 9; Length 547;  
 Best Local Similarity 86.1%; Pred. No. 1,4e-77;  
 Matches 470; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 219 GTGGTTGAGACACACACAGTGGAGCTGCGCCACACTGTAGATCCAGGTGGCAAAA 278  
 Db 2 GTGGTTGAGACACACAGTGGAGCTGCGCCACACTGTAGATCCAGGTGGCAAAA 61  
 QY 279 GACATATTATAACAACTTTCTTGTGACTGCGCCAGAGAGAGAGATGCTTGTGATGC 338  
 Db 62 AACCATTTATCAATTAAGCTCTTCTTGTGACTGCGCCAGAGAGAGAGATGCTTGTGATGC 121  
 QY 339 AGAATTTCTTAAGAAATGAACTGACAGCGTCAAGCTCAAGCTTCCAGAAAGACAGGA 398  
 Db 122 AGAATTTCTTAAGAAATGAACTGATTAATACCAGAGCCCTGCTTCCAGAAAGAGAGA 181  
 QY 399 GAAAGGAGACAGCCAGGCTATATGACACTCTACGAGGACACCTGGAAAGACCAATGC 458  
 Db 182 AAAAGGAGACAGCCAGGCTATATGACACTCTACGAGGACACCTGGAAAGACCAATGC 241  
 QY 459 TACCGTGAAGTCCCTACAGAAAGCCCTTAACCAAGCAGAGATGCTGTTCACACCTGAA 518  
 Db 242 CACGTGTGAATCCCTGCAAGAGGCTTATAGACAGCCGAAATGCTGTGTCTCACACCTCAA 301  
 QY 519 AAAACAGATGAAGTTCTGAGAGCGGACAGGAGATGAGACCAACAAAGCTGGAGAGGC 578  
 Db 302 GAAACAGATGAAGTTCTGAGAGCGGACAGGAGATGAGACCAACAAAGCTGGAGAGGC 361  
 QY 579 CCACGACTCAAGTCAAGATGAAGAAACCATGAGCAAAATGAGTCCCTACTCCAGAGCCA 638  
 Db 362 CGCCGAGCTCGGAGCAAGATGAAGAACTATGAGACGAGATGAGTCTCTACTCCAGAGCCA 421  
 QY 639 GCGTCTGAGGTGAGAGAGATGATTCGACATGAGTGTGAGACAGTCAAGCGGTGAGCA 698  
 Db 422 GCGGCTGAGGTGAGAGAGATGATTCGACATGAGTGTGAGACAGTCAAGCGGTGAGCA 481  
 QY 699 GCTGCTGTGTACTGCGTGTCTCCCAAGAAAGATGAGATCTGAGAGCTGGGAA 758  
 Db 482 GCTGCTGTGTACTGCGTGTCTCCCAAGAAAGATGAGATCTGAGAGCTGGGAA 541  
 QY 759 GCGCCAC 764  
 Db 542 GGCTTC 547

RESULT 30  
 BE031869 450 bp mRNA linear EST 19-NOV-2001  
 LOCUS  
 DEFINITION BB031869 RIKEN full-length enriched, mammary gland RCB-0527  
 Jy9-MC(B) cDNA Mus musculus cDNA clone G930019H09 3', mRNA  
 sequence.  
 ACCESSION BB031869  
 VERSION BB031869.1 GI:17010112  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 450)  
 AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,  
 Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii,  
 Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T.,  
 Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
 Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,  
 Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,  
 A., Takahashi,F., Takaku-Akenita,S., Tanaka,F., Tomaru,A., Toyota,T.,  
 Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
 2001)

JOURNAL  
COMMENT

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

## FEATURES

Location/Qualifiers

1. 450

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="G930019B09"

/clone\_1lb="RIKEN full-length enriched, mammary gland

RCB-0527 Jyg-MC(B) cDNA"

/tissue\_type="mammary gland"

/cell\_line="RCB-0527 Jyg-MC(B)"

BASE COUNT  
ORIGIN

105 a 102 c 114 g 129 t

Query Match

21.4%; Score 422; DB 9; Length 450;

Best Local Similarity 97.6%; Pred. No. 4.6e-77;

Matches 439; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

OY 1507 ACCTTCTATGTCACTGAACGGTGCACAGATGATGTTGCATTAATGATGGCCAGACCT 1566

Db 1 ACCCTTTATGTCACTGAACGGTGCACAGATGATGTTGCATTAATGATGGCCAGACCT 60

OY 1567 GGGTAAACGGAGATGTTTGGAAAGATGGCTCCTTGGACACAGCCAAAGAGATGGCC 1626

Db 61 GGGTAAACGGAGATGTTTGGAAAGATGGCTCCTTGGACACAGCCAAAGAGATGGCC 120

OY 1627 AGAAACACACTTCTGTGTCTACATGCGCCCTGCACACACTGGAAGCCACATGACAG 1686

Db 121 AGAAACACACTTCTGTGTCTACATGCGCCCTGCACACACTGGAAGCCACATGACAG 180

OY 1687 TTTTACTGTTCCGATCAGCAGGGCTACTTCCAGTTGACAGGGTTTGTATAGCTAAC 1746

Db 181 TTTTACTGTTCCGATCAGCAGGGCTACTTCCAGTTGACAGGGTTTGTATAGCTAAC 240

OY 1747 CAGGTGTGGGTGGAGTCTCTTTTGTATAGAACAGGGTGCACATGACTTAATGAT 1806

Db 241 CAGGTGTGGGTGGAGTCTCTTTTGTATAGAACAGGGTGCACATGACTTAATGAT 299

OY 1807 GGAAGTGTGGAGATCTTATGACAGGTGAGAGACCTGCGCTTAATGCTGCTGCT 1866

Db 300 GGAAGTGTGGAGATCTTATGACAGGTGAGAGACCTGCGCTTAATGCTGCTGCT 359

OY 1867 CCAGCTTATGCTTGAATTAATGGGTGAGGTGATAGGAAAGTTGGGGAAGTTT 1926

|||||

Db 360 CCAGCTTATGCTTGAATTAATGGGTGAGGTGATAGGAAAGTTGGGGAAGTTT 419

OY 1927 CTGTCTAAATAAAGGAGCTTTCTTC 1956

Db 420 CTGTCTAAATAAAGGAGCTTTCTTC 449

RESULT 31

BB830067 453 bp mRNA linear EST 19-NOV-2001

LOCUS BB830067 RIKEN full-length enriched, mammary gland RCB-0527

DEFINITION Jyg-MC(B) cDNA Mus musculus cDNA clone G930010B01 3', mRNA

sequence.

ACCESSION BB830067

VERSION BB830067.1 GI:17008310

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

Location/Qualifiers

1. 453

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="G930010B01"

/clone\_1lb="RIKEN full-length enriched, mammary gland

RCB-0527 Jyg-MC(B) cDNA"

/tissue\_type="mammary gland"

/cell\_line="RCB-0527 Jyg-MC(B)"

BASE COUNT  
ORIGIN

104 a 99 c 119 g 131 t

Query Match 20.6%; Score 407.8; DB 9; Length 453;  
Best Local Similarity 97.8%; Pred. No. 4e-74;  
Matches 445; Conservative 0; Mismatches 7; Indels 3; Gaps 3;

QY 1502 TGGATACCTTCTTATGTCAGTGAACGGTGAACAGATGATGTTGCAATTAGTGGCCAA 1561  
1 TGGATACCTTCTTATGTCAGTGAACGGTGAACAGATGATGTTGCAATTAGTGGCCAA 59  
DB 1562 GACCTGGCTAACCGAGAGTCTTTTGAAGATGGCTCCTTGACCACTCCAGAGAGA 1621  
60 GACCTGGCTAACCGAGAGTCTTTTGAAGATGGCTCCTTGACCACTCCAGAGAGA 118  
QY 1622 TGGCCAGAAAAACACACTTCTGTTGCTACTGCGCCCTGCACACACATGGGAAGCCATG 1681  
119 TGGCCAGAAAAACACACTTCTGTTGCTACTGCGCCCTGCACACACACT-GGAGGCCAATG 177  
QY 1682 ACCAGTTTACTGTTCCGATCAGCAGGGCCATCTTCCAGTTGACAGGTTTGTCTATAGCT 1741  
178 ACCAGTTTACTGTTCCGATCAGCAGGGCCATCTTCCAGTTGACAGGTTTGTCTATAGCT 237  
QY 1742 ACAACGAGGTGGCTGGACCTCTTGTGTTTATAGACAGGCTCATTTGACTCTAG 1801  
238 ACAACGAGGTGGCTGGACCTCTTGTGTTTATAGACAGGCTCATTTGACTCTAG 297  
QY 1802 TGGATGGAGTGTGTGAGATCTTATGTCAGGTCGAGACCTGCGCTTGAACCTGCGC 1861  
298 TGGATGGAGTGTGTGAGATCTTATGTCAGGTCGAGACCTGCGCTTGAACCTGCGC 357  
QY 1862 TGGCTCCAGCTTATGCTTGAATTAATGAGGTGAGTGTGATAGGAAAGTTGGGAA 1921  
358 TGGCTCCAGCTTATGCTTGAATTAATGAGGTGAGTGTGATAGGAAAGTTGGGAA 417  
QY 1922 GTTTTGTGTGAATAATAAAGGATCTTTCTTC 1956  
418 GTTTTGTGTGAATAATAAAGGATCTTTCTTC 452

RESULT 32  
BB835693 412 bp mRNA linear EST 19-NOV-2001  
LOCUS BB835693 RIKEN full-length enriched, mammary gland RCB-0527  
DEFINITION Jyg-MC(B) CDNA Mus musculus cDNA clone G930038G09 3', mRNA  
sequence.  
ACCESSION BB835693  
VERSION BB835693.1 GI:17013936  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 412)  
REFERENCE 1 (bases 1 to 412)  
AUTHORS Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanaoka,T.,  
Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imotani,K., Ishii  
,Y., Ito,M., Kawai,J., Kojima,Y., Kondo,H., Kouda,M., Matsuyama,T.,  
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,  
Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,  
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa  
,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,  
, Watabiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
, RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.  
2001)

TITLE Unpublished (2001)  
JOURNAL Contact: Yoshihide Hayashizaki  
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sphiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-resgsc.riken.go.jp,  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh

Query Match 20.6%; Score 406.2; DB 9; Length 412;  
Best Local Similarity 99.3%; Pred. No. 8.6e-74;  
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1546 GCAATTAGTGGGCGCAGACCTGCTTAACCGAAGTGTGTTTGAAGATGCTCCTTGG 1605  
1 GCATTTAGTGGGCGCAGACCTGCTTAACCGAAGTGTGTTTGAAGATGCTCCTTGG 60  
QY 1606 ACCAGTCCAGAGAGATGCCAGAAACACACTTCTGTTTCACTGCGCCCTGCACAC 1665  
61 ACCAGTCCAGAGAGATGCCAGAAACACACTTCTGTTTCACTGCGCCCTGCACAC 120  
QY 1666 ACTGGGAAGCCACATGACACAGTTACTGTTCCATGACGAGGCCCTCACTCCAGT 1725  
121 ACTGGGAAGCCACATGACACAGTTACTGTTCCATGACGAGGCCCTCACTCCAGT 180  
QY 1726 GGTTCCTTATAGCTCAACACAGGTGGCTGACCTCTTGTGTTTATAGAAGAGG 1785  
181 GGTTCCTTATAGCTCAACACAGGTGGCTGACCTCTTGTGTTTATAGAAGAGG 240  
QY 1786 TCACATTGACTCTAAGTGAATGAGGAGTGGAGATCTTATGACGAGCTGAGACCTG 1845  
241 TCACATTGACTCTAAGTGAATGAGGAGTGGAGATCTTATGACGAGCTGAGACCTG 300  
QY 1846 CGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905  
301 CGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 1906 GGAAGAGTGGGAGAGTCTGTTGAATAAAGGATCTTTCTTC 1956  
361 GGAAGAGTGGGAGAGTCTGTTGAATAAAGGATCTTTCTTC 411

FEATURES  
source location/Qualifiers  
1..412  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_lib="G930038G09"  
/clone\_lib="RIKEN full-length enriched, mammary gland  
RCB-0527 Jyg-MC(B) CDNA"  
/tissue\_type="mammary gland"  
/cell\_line="RCB-0527 Jyg-MC(B)"  
BASE COUNT 96 a 91 c 111 g 114 t  
ORIGIN

Query Match 20.6%; Score 406.2; DB 9; Length 412;  
Best Local Similarity 99.3%; Pred. No. 8.6e-74;  
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1546 GCAATTAGTGGGCGCAGACCTGCTTAACCGAAGTGTGTTTGAAGATGCTCCTTGG 1605  
1 GCATTTAGTGGGCGCAGACCTGCTTAACCGAAGTGTGTTTGAAGATGCTCCTTGG 60  
QY 1606 ACCAGTCCAGAGAGATGCCAGAAACACACTTCTGTTTCACTGCGCCCTGCACAC 1665  
61 ACCAGTCCAGAGAGATGCCAGAAACACACTTCTGTTTCACTGCGCCCTGCACAC 120  
QY 1666 ACTGGGAAGCCACATGACACAGTTACTGTTCCATGACGAGGCCCTCACTCCAGT 1725  
121 ACTGGGAAGCCACATGACACAGTTACTGTTCCATGACGAGGCCCTCACTCCAGT 180  
QY 1726 GGTTCCTTATAGCTCAACACAGGTGGCTGACCTCTTGTGTTTATAGAAGAGG 1785  
181 GGTTCCTTATAGCTCAACACAGGTGGCTGACCTCTTGTGTTTATAGAAGAGG 240  
QY 1786 TCACATTGACTCTAAGTGAATGAGGAGTGGAGATCTTATGACGAGCTGAGACCTG 1845  
241 TCACATTGACTCTAAGTGAATGAGGAGTGGAGATCTTATGACGAGCTGAGACCTG 300  
QY 1846 CGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1905  
301 CGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 1906 GGAAGAGTGGGAGAGTCTGTTGAATAAAGGATCTTTCTTC 1956  
361 GGAAGAGTGGGAGAGTCTGTTGAATAAAGGATCTTTCTTC 411

RESULT 33  
BB777900 423 bp mRNA linear EST 15-NOV-2001  
LOCUS BB777900 RIKEN full-length enriched, nullipotent stem cell CRL-2070  
DEFINITION NE CDNA Mus musculus cDNA clone G430042L20 3', mRNA sequence.  
ACCESSION BB777900  
VERSION BB777900.1 GI:16938600  
KEYWORDS EST.

SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae: Mus.  
AUTHORS 1 (bases 1 to 423)  
Akimura,T., Aikawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T., Imoto,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Kono,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toyota,T., Watanishi,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
Riken Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001).  
TITLE Unpublished (2001)  
JOURNAL Contact: Yoshihide Hayashizaki  
COMMENT Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagii,K., Fujiwaka,S., Inoue,K., Togawa,Y., Ikawa,M., Ohara,E., Watanishi,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.  
e mouse tissues.  
FEATURES  
LOCATION/Qualifiers  
1..423  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="G430042L20"  
/clone\_lib="RIKEN full-length enriched, Nullipotent stem cell CRT-2070 NE cDNA"  
/cell\_type="Nullipotent stem cell"  
/cell\_line="CRL-2070 NE"  
BASE COUNT 100 a 93 c 111 g 119 t  
ORIGIN  
Query Match 20.3%; Score 401.4; DB 9; Length 423;  
Best Local Similarity 96.3%; Pred. No. 8.4e-73;  
Matches 416; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 1534 AGAGGATGTTGCAATAGTGGGCAAGACCTGCTAACCGGAAGTGTGGGAAGAT 1593  
|||||  
DB 1 AGAGCATGTTGCAATAGTGGGCAAGACCTGCTAACCGGAAGTGTGGGAAGAT 59  
QY 1594 GGTCTCTTGGACAGTCCAGAGAGATGCCAGAAAAACACACTCTCTGTGTTCATGC 1653  
|||||  
DB 60 GATCTCTTGGACAGTCCAGAGAGATGCCAGAAAAACACACTCTCTGTGTTCATGC 119  
QY 1654 GCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATAGCAGGGCCCTAC 1713  
|||||  
DB 120 GCCCTGCACACACTGGGAAGCCACATGACAGTTTACTGTTCCGATAGCAGGGCCCTAC 179

QY 1714 TTCAGTTGCAGGTTTTGCTTATAGTACACAGGAGTGTGGCTGACTCCCTTTGTTTT 1773  
|||||  
DB 180 TTCAGTTGCAGGTTTTGCTTATAGTACACAGGAGTGTGGCTGACTCCCTTTGTTTT 239  
QY 1774 TATAGACAGGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1833  
|||||  
DB 240 TATAGACAGGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 299  
QY 1834 TGGAGAGCCCTGCGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1893  
|||||  
DB 300 TGGAGAGCCCTGCGCTTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 359  
QY 1894 GAGTGTGATAGGGAAGGTTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGG 1953  
|||||  
DB 360 GAGTGTGATAGGGAAGGTTGGGAGTGTGGGAGTGTGGGAGTGTGGGAGTGTGGG 419  
QY 1954 TTC 1956  
|||  
DB 420 TTC 422  
RESULT 34  
A1428513 403 bp mRNA linear EST 09-MAR-1999  
LOCUS  
DEFINITION mo37h05.x1 life Tech mouse embryo 13 5dpc 10666014 Mus musculus  
ACCESSION A1428513  
VERSION A1428513.1 GI:4274439  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae: Mus. 1 (bases 1 to 403)  
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harrey,N., Schurk,R., Rittler,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.  
The WashU-NCI Mouse EST Project 1999  
JOURNAL Unpublished (1999)  
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
This clone was previously sequenced on the 5' end only, this new data is from the 3' end  
Possible reversed clone: polyT not found  
High quality sequence stop: 402.  
FEATURES  
LOCATION/Qualifiers  
1..403  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:555801"  
/clone\_lib="Life Tech mouse embryo 13 5dpc 10666014"  
/tissue\_type="embryo"  
/dev\_stage="13 5dpc embryos"  
/lab\_host="DH10B"  
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site\_1: Salt; Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 13.5dpc embryos. pCMV-SPORT2 vector."  
BASE COUNT 110 a 108 c 88 g 97 t  
ORIGIN  
Query Match 20.2%; Score 399.8; DB 9; Length 403;  
Best Local Similarity 99.5%; Pred. No. 1.8e-72;

[illegible]

RESULT	35
AM823060/c	
LOCUS	AM823060
DEFINITION	ur79c12.x1 NCI-CGAP_Mam6 Mus musculus CDNA clone IMAGE:3156502 3', mRNA sequence.
	391 bp mRNA linear EST 17-MAY-2000

ACCESSION	AW823060	
VERSION	AW823060.1	GI:7916137
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus  
TITLE 1 (pases 1 to 391)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>  
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),  
Tumor Gene Index  
unpublished (1997)  
Other-ESTs: ur79c12.y1

Contact: Robert Strausberg, Ph.D.  
 Email: cga@bs-r@mail.nih.gov  
 Tissue Procurement: Jeffrey Green M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
[image.llnl.gov/image/html/tiresources.shtml](http://image.llnl.gov/image/html/tiresources.shtml)  
 MGI:1059258  
 Seq primer: -40UP from Gidco  
 'High quality sequence stop: 380.  
 Location/Qualifiers  
 I. .391

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3156502"
/clone_lib="NCI_CGP_Mam6"
/sex="female, virgin"
/tissue_type="infiltrating ductal carcinoma"
/dev_stage="5 months"
/lab_host="DH10B"

```

BASE COUNT	107	a	106	c	87	g	90	t	1	others
ORIGIN	/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Jeffrey Green, M.D., NIH"									

Query Match	19.7%	Score 388.4;	DB 9;	Length 391;
Best Local Similarity	99.5%;	Pred. No. 4.1e-70;		
Matches 389;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0

Oy	1563	ACCTGGCTAACCGGAAGTGTGTTTGGAAGAATGGCTCCTCTGGACCAGTCACAAGAGAGAT	1622
Db	391	ACCTGGCTAACCGGAAGTGTGTTTGGAAGAATGGCTCCTCTGGACCAGTCACAAGAGAGAT	332
Oy	1623	GCCCAAAAACACACTTCCCTGTTCACACGCCCCCTGCACCACTAGGGAAAGCCACATTA	1682
Db	331	GCCCAAAAACACACTTCCCTGTTCACACGCCCCCTGCACCACTAGGGAAAGCCACATTA	272
Oy	1683	CCAGTTTACTGTTCCGATCAGCAGGGCCACTTCCAGTTGCAGGGTTTTGCTTATAGCTA	1742
Db	271	CCAGTTTACTGTTCCGATCAGCAGGGCCACTTCCAGTTGCAGGGTTTTGCTTATAGCTA	212
Oy	1743	CACCAAGGTGTGGCTGGACTCCTTTTGTATTATAGAACAGGGTCACTTGACTCTAACT	1802
Db	211	CACCAAGGTGTGGCTGGACTCCTTTTGTATTATAGAACAGGGTCACTTGACTCTAACT	152
Oy	1803	GGATGGGAATGCTGAGAGATCCTATGACAGGCTGGAGAGACCCTGCGTTGAATCCTCGCCT	1862
Db	151	GGATGGGAATGCTGAGAGATCCTATGACAGGCTGGAGAGACCCTGCGTTGAATCCTCGCCT	92
Oy	1863	GGCTCAGCTTTTGCTTAAATTTATNGGGGTGAGGTGGATAGGGAAGTTGGGGGAAG	1922
Db	91	GGCTCAGCTTTTGCTTAAATTTATNGGGGTGAGGTGGATAGGGAAGTTGGGGGAAG	32
Oy	1923	TTTTCTGTAAAAATAAAGGAGCTTTTTC	1953
Db	31	TTTTCTGTAAAAATAAAGGAGCTTTTTC	1

RESULT	36		
BB837528			
LOCUS		393 bp	mRNA
DEFINITION	BB837528	RIKEN full-length enriched, mammary gland	EST 19-NOV-2000
	Jy9-MC(B)	CDNA Mus musculus clone G930047020 3'	mRNA
		sequence.	
ACCESSION	BB837528		
VERSION	BB837528.1	GI:17015771	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		

Euarchyotia; Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia: Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 393)

Akimura, T., Akakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Imochani, K., Ishi, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Nunaseki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sugabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akibira, S., Tanaka, T., Tomaru, A., Toyota, R., Watahiki, A., Yasunishi, A., Yamatsutsu, M. and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22, Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,  
URL: http://genome-gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
Wegli, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura  
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multichipillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (http://genome-gsc.riken.go.jp) for  
further details.  
e mouse tissues.  
Location/Qualifiers

FEATURES  
Source  
1. 393  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="G93004.7020"  
/clone\_lib="RIKEN full-length enriched, mammary gland  
RCB-0527 Jyg-MC(B) cDNA"  
/tissue\_type="mammary gland"  
/cell\_line="RCB-0527 Jyg-MC(B)"  
BASE COUNT 88 a 87 c 106 g 112 t  
ORIGIN

Query Match 19.5%; Score 385.6; DB 9; Length 393;  
Best Local Similarity 99.0%; Pred. No. 1.6e-69;  
Matches 388; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1555 CTGGTACCGGAATGTTTGGAGAATGGCTCTCTTGACACAGTCCAGAGAGATGC 1624  
Db 1 CTGGCTAACCGGAATGTTTGGAGAATGGCTCTCTTGACACAGTCCAGAGAGATGC 60  
Oy 1625 CCAGAAACACACTCTCTGTGTCTACCTGCGCCCTGCACACACAGGAGGACATGACC 1684  
Db 61 CCAGAAACACACTCTCTGTGTCTACCTGCGCCCTGCACACACAGGAGGACATGACC 120  
Oy 1685 AGTTACTGTTCCGATCAGACAGGGCCCTACTTCCAGTTGACAGGGTTTGTCTATAGCTACA 1744  
Db 121 AGTTACTGTTCCGATCAGACAGGGCCCTACTTCCAGTTGACAGGGTTTGTCTATAGCTACA 180  
Oy 1745 ACCAGGTGGCTGAGACCTCTTTTATATAGACAGGGTGCACATTGACTTAAGTGG 1804  
Db 181 ACCAGGTGGCTGAGACCTCTTTTATATAGACAGGGTGCACATTGACTTAAGTGG 240  
Oy 1805 ATGGAGTGTGAGAGATCTATGACAGCTGGAGGACCTGGCTGTAAGTCTGCTGC 1864  
Db 241 ATGGAGTGTGAGAGATCTATGACAGCTGGAGGACCTGGCTGTAAGTCTGCTGC 300  
Oy 1865 CTCGAGCTTATGCTTGAATATGAGGTGAGTGTGATAGGAAAGTTGGGGAAGTT 1924  
Db 301 CTCGAGCTTATGCTTGAATATGAGGTGAGTGTGATAGGAAAGTTGGGGAAGTT 360  
Oy 1925 TTCTGTGTAATAAATAAAGGATCTTTCTTC 1956  
Db 361 TTCTGTGTAATAAATAAAGGATCTTTCTTC 392

RESULT 37  
BI341090 571 bp mRNA linear EST 30-JUL-2001  
LOCUS BI341090  
DEFINITION 368355 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BI341090

VERSION BI341090.1 GI:15034379  
KEYWORDS EST.  
SOURCE Sus scrofa  
ORGANISM pig.  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE  
1 (bases 1 to 571)  
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,  
and Keeler, J.W.  
Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
Unpublished (2000)  
Contact: Smith TPL

TITLE  
JOURNAL  
COMMENT  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@meat.ars.usda.gov  
Single pass sequencing. Bases called and alt-trimmed with phred  
v0.980904.e. Vector identified by cross-match with the -minscore 18  
and -mismatch 12 options.  
PCR primers  
FORWARD: AGGAACACGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCACGACG  
Plate: 107 row: K column: 10  
Seq primer: ATTTAGCTGACCTATGAC.

FEATURES  
Source  
1. 571  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: PCMV SPORT6; site\_1: XbaI; site\_2: XhoI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."

BASE COUNT 154 a 162 c 146 g 109 t  
ORIGIN

Query Match 18.9%; Score 373.6; DB 10; Length 571;  
Best Local Similarity 84.8%; Pred. No. 4.5e-67;  
Matches 431; Conservative 0; Mismatches 74; Indels 3; Gaps 1;

Oy 81 TGTTCCTCGGCTGCTGAGTCGAGCCATATGCTATCTCTCTGCTGCTACTATCTG 140  
Db 66 TGGTACCTGGGCACTTGAGACAG--TCATGCTATGCTGCTGCTGCTGCTATATG 122  
Oy 141 CTCCGACTTTCGATCACTCCCGTGCAGCTGGCCATCCAGTGGGCCACTTTTCA 200  
Db 123 CTCCGACTTTCGATCACTCCCGTGCAGCTGGCCATCCAGTGGGCCACTTTTCA 182  
Oy 201 TCTCAATGCTTAATCCAGTGGTTGAGACAGCAAGTGGAGCTGCCACAGTGTAG 260  
Db 183 CTTCGAGTCTTAATTCAGTGGTTGAGACAGCAAGTGGAGCTGCCACAGTGTAG 242  
Oy 261 AATCCAGTTGGCAAAAGACTATTATTAACAACCTTTTCTTGACCTGCGCCAGAGA 320  
Db 243 AATCCAGTTGGCAAAAGACTATTATTAACAACCTTTTCTTGACCTGCGCCAGAGA 302  
Oy 321 GGAAGATGCTTGAATGAGATTCCTTAAGAATGAAGTGAACAGCGCTCAAGTCACT 380  
Db 303 GGAAGATGCTTGAATGAGATTCCTTAAGAATGAAGTGAATGAATCAAGCGCTGCT 362  
Oy 381 TTCCAGAAAGACAGGAGAAAGGAGCAGCAGCCATTTCGACACTTACGGGACAC 440  
Db 363 TTCCAGAAAGACAGGAGAAAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 422  
Oy 441 CTGGAAGACGCAATGCTTACCTGAGTCCCTTACAGAGCCCTTAAACAAGCAGAGAT 500  
Db 423 GCTGAAGAGCGCAACCTGCTGTAATCCCTGACAGAGCCCTTAAACAAGCAGAGAT 482

```

QY 501 GCTGTGTCACCCGTGAAGAAACAGATGATTCCTGAGCAGCGCAGATGAGACCA 560
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 483 GCTGTGTCACCCGTGAAGAAACAGATGATTCCTGAGCAGCGCAGATGAGACCA 542
QY 561 ACAAGCTGGAGAGAGCCGACGCTC 588
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 ACAAGCAGGAGAGAGCCGCGCAGCTC 570

RESULT 38
BF038722
LOCUS 769 bp mRNA linear EST 20-OCT-2000
DEFINITION 60146064f1 NIH_MGC_66 Homo sapiens cDNA clone IMAGE:3864082 5',
            mRNA sequence.
ACCESSION BF038722
VERSION BF038722.1 GI:10745770
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
            Tissue Procurement: DCTD/DRP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LAM9605 row: a column: 11
            High quality sequence stop: 632.
            Location/Qualifiers
                1. 769
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:3864082"
                /clone_1id="NIH_MGC_66"
                /tissue_type="adenocarcinoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: ovary; Vector: pCMV-SPORT6; Site:1; NotI;
                Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.8 kb. Library constructed by Life
                Technologies."
BASE COUNT 184 a 212 c 217 g 156 t
ORIGIN

Query Match 18.6%; Score 367.6; DB 10; Length 769;
Best Local Similarity 79.0%; Pred. No. 7.4e-66;
Matches 463; Conservative 0; Mismatches 119; Indels 4; Gaps 2;

QY 940 AACCTTGAGCGCTGCTC-GGCACCAATGAGAGCGTCAGCGCGCTTTTGGAGGCC 998
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 AACCTTGAGCGCTGCTCAGCAGCGCTGAGAGCTGTCAGCGCGCTTTTGGAGGCC 60
QY 999 AGCCCTGTGAGATGATGAACCCGAGGCTTCACAGCCACCTTCGCTGATGAGATGA 1058
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 AGCCCTGTGGA--GGGATGTCGAAAGCTCCGCGGCGCATCTTCGCTGATGATGA 117
QY 1059 TCTCAATACCACTTTGATGAATACCCCTCAACCCAGACCTCTGCTCCACATGG 1118
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 118 TCTCAATGCTACTTTGATGATGATACCCCGCCGCGCTTCACCTCCACAGATGG 177
QY 1119 CCTGCCAAGAGCTGCTGAGAGGAGCGCTCCCAATGAGATGTCCTCAAGAA 1178
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 TTACTACAGAAACTTTCCTAGAGAGATCACTCCCAATTCAGATGTCGCCAAGAA 237
QY 1179 GGTGCAAAAGTCTCCAGCGGAGTCCAGCTCTCACTGGGTGCCAGCATGTGTAG 1238

```

```

Db 238 GATATGCAAGAGGCCCGCAGAGAGAGTCCAGCTCTCACTGGTGCCAGAGCTGTG 297
QY 1239 AGAGCTGATGAGAGAACTGCTGTGCTTCCTCCCTCTTCATCCGGAATGTCCTGG 1298
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 298 AGAGCTGATGAGAGAACTGCTGTGCTTCCTCCCTCTTCATCCGGAATGTCCTAG 357
QY 1299 TCAGAAACAGCCCGCAGACACAGCAATCCCGAGAGCAGCAGATGATGTAAGAT 1358
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 358 CCAGAAACAGCCCGCAGAGGCCAGCTGAGTCTCTTCGACAGAAATGATGTAAGAG 417
QY 1359 AGCTTTGATGAGCTTGGAGAGAGACAAATTCATCCAGCTTGGAGACACATAT 1418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 AGCTTGTGATGGGCTCGCTGCGCCGACAAATTCATCCAGCTTGTGACAGATCAT 477
QY 1419 CCGACCAATGCGCTTGAATCCAGGCCAGAGTAAGACAGAAATGAAATGAAGCTG 1478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 478 CCGCCCATGCTGTGTAAGCCCAAGACCAAGTTAAGCAGAGGTGAGGGTGAAGACAGT 537
QY 1479 GAGTTCGCTCCCGCAGCCCAAGCTGATACCTCTTATGTCAGTGA 1524
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 538 GCTTCTCTCTCCAGGCCAAGCTGAGACCTTCTGTCGTCTGA 583

RESULT 39
AL529568
LOCUS 932 bp mRNA linear EST 13-FEB-2001
DEFINITION AL529568 LTL_NFL001_NBC4 Homo sapiens cDNA clone CSDD006YH09 5
            prime, mRNA sequence.
ACCESSION AL529568
VERSION AL529568.1 GI:12793061
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
JOURNAL Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
            Location/Qualifiers
                1. 932
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="CSDD006YH09"
                /clone_1id="LTL_NFL001_NBC4"
                /sex="male"
                /tissue_type="neuroblastoma cells"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed
                by Life Technologies. Contact: Feng Liang Life
                Technologies, a division of Invitrogen 9800 Medical Center
                Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                8371 Email : fliang@lifetech.com URL :
                http://fulllength.invitrogen.com"
BASE COUNT 226 a 253 c 242 g 200 t 11 others
ORIGIN

Query Match 18.0%; Score 354.8; DB 9; Length 932;
Best Local Similarity 75.9%; Pred. No. 3.2e-63;
Matches 481; Conservative 7; Mismatches 140; Indels 6; Gaps 4;

QY 769 GAAGTGGCTGACAGGTTGAGAGAGATTTGCTCTTAGAGACCAAGTGAAGACTCTG 828
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



